PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/435, 14/00, 7/06, C12N 5/10, 15/10, 15/11, 15/12, 15/63, C12Q 1/68, A61K 38/16

A1

(11) International Publication Number: **WO 00/09551**

(43) International Publication Date: 24 February 2000 (24.02.00)

(21) International Application Number: PCT/US99/18117

(22) International Filing Date: 10 August 1999 (10.08.99)

(30) Priority Data:

60/095,880	10 August 1998 (10.08.98)	US
09/306,111	6 May 1999 (06.05.99)	US

(71) Applicant: GENETICS INSTITUTE, INC. [-/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Street, Newton, MA 02160 (US). McCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LaVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). COLLINS-RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 12 Foxrock Court, Dublin 18 (IE). DiBLASIO-SMITH, Elizabeth; 17 Chestnut Road, Tyngsboro, MA 01879 (US). WIDOM, Angela; 19 Cherokee Road, Acton, MA 01720 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: HUMAN CHORDIN-RELATED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel human chordin-related proteins and polynucleotides encoding them are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN CHORDIN-RELATED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 09/306,111, filed May 6, 1999, which is a continuation-in-part of application Ser. No. 60/095,880, filed August 10, 1998.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. Specifically, the present invention relates to a novel family of purified proteins designated chordin-related proteins, DNA encoding them, and processes for obtaining them. These proteins may be used to induce and/or regulate bone and/or cartilage or other connective tissue formation, and in wound healing and tissue repair. These proteins may also be used for augmenting the activity of other bone morphogenetic proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

The search for the molecule or molecules responsible for the bone-, cartilage-, and other connective tissue-inductive activity present in bone and other tissue extracts has led to the discovery and identification of a several groups of molecules, such as the Bone Morphogenetic Proteins (BMPs). The unique inductive activities of these proteins, along with their presence in bone, suggests that they are important regulators of bone repair processes, and may be involved in the normal maintenance of bone tissue. There is a need to identify whether additional proteins, particularly human proteins, exist which play a role in these processes. *Xenopus chordin* is a molecule which contributes to dorsoventral patterning by binding to BMP-4. (See Piccolo *et al.*, Cell, 86:589-98 (1996).) The nucleotide and amino acid sequences of *xenopus chordin* are described in Lasai et al., Cell, 79:779-790 (1994). The *xenopus chordin* gene has been described as being expressed in the frog embryo head, trunk, and tail organizer regions during gastrulation, and as being capable of inducing secondary axes in frog embryos, and rescuing axis formation in ventralized frog, as well as modifying mesoderm induction. *Ibid.* In addition, *xenopus chordin* has been shown to induce anterior neural markers in the absence of mesoderm induction. Sashai et al., Nature, 376:333-336 (1995). Human chordin is the human homolog of *xenopus chordin*, and is described, for example, in LaVallie *et al.*, 1998, U.S. Patent No. 5,846,700 . The present invention relates to the identification of novel human proteins with structures

and activities similar to chordin proteins, which the inventors have designated human chordin-related proteins.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 157 to nucleotide 1356;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dj167_2 deposited under accession number ATCC 98818;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

(e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dj167_2 deposited under accession number ATCC 98818;

(f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:2;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ;

(k) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h); and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h) and that has a length that is at least 25% of the length of SEQ ID NO:1.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 157 to nucleotide 1356; the nucleotide sequence of the full-length protein coding sequence of clone dj167_2 deposited under accession number ATCC 98818; or the nucleotide sequence of a mature protein coding sequence of clone dj167_2 deposited under accession number ATCC 98818. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:2, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 195 to amino acid 204 of SEQ ID NO:2.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:1, but excluding the poly(A) tail at the 3' end of SEQ ID NO:1; and

(ab) the nucleotide sequence of the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

(ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:1, but excluding the poly(A) tail at the 3' end of SEQ ID NO:1; and

(bb) the nucleotide sequence of the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:1, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:1 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:1 , but excluding the poly(A) tail at the 3' end of SEQ ID NO:1. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:1 from nucleotide 157 to nucleotide 1356, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:1 from nucleotide 157 to nucleotide 1356, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:1 from nucleotide 157 to nucleotide 1356.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) a fragment of the amino acid sequence of SEQ ID NO:2, the fragment comprising eight contiguous amino acids of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably

comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:2, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 195 to amino acid 204 of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1383 to nucleotide 4490;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1485 to nucleotide 4490;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 3645 to nucleotide 4343;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dj167_19 deposited under accession number ATCC 207090;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dj167_19 deposited under accession number ATCC 207090;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:4;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ;

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j); and

(n) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j) and that has a length that is at least 25% of the length of SEQ ID NO:3.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 1383 to nucleotide 4490; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1485 to nucleotide 4490; the nucleotide sequence of SEQ ID NO:3 from nucleotide 3645 to nucleotide 4343; the nucleotide sequence of the full-length protein coding sequence of clone dj167_19 deposited under accession number ATCC 207090; or the nucleotide sequence of a mature protein coding sequence of clone dj167_19 deposited under accession number ATCC 207090. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 637 to amino acid 1036. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:4, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 513 to amino acid 522 of SEQ ID NO:4.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:3, but excluding the poly(A) tail at the 3' end of SEQ ID NO:3; and

(ab) the nucleotide sequence of the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

(ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:3, but excluding the poly(A) tail at the 3' end of SEQ ID NO:3; and

(bb) the nucleotide sequence of the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:3 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:3 , but excluding the poly(A) tail at the 3' end of SEQ ID NO:3. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3 from nucleotide 1383 to nucleotide 4490, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:3 from nucleotide 1383 to nucleotide 4490, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:3 from nucleotide 1383 to nucleotide 4490. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3 from nucleotide 1485 to nucleotide 4490, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:3 from nucleotide 1485 to nucleotide 4490, to a nucleotide sequence corresponding to the 3' end

of said sequence of SEQ ID NO:3 from nucleotide 1485 to nucleotide 4490. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3 from nucleotide 3645 to nucleotide 4343, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:3 from nucleotide 3645 to nucleotide 4343, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:3 from nucleotide 3645 to nucleotide 4343.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 637 to amino acid 1036;

(c) a fragment of the amino acid sequence of SEQ ID NO:4, the fragment comprising eight contiguous amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 637 to amino acid 1036. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:4, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 513 to amino acid 522 of SEQ ID NO:4.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 71 to nucleotide 1441;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 1441;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dw665_4 deposited under accession number ATCC 98818;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dw665_4 deposited under accession number ATCC 98818;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ;

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:5.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 71 to nucleotide 1441; the nucleotide sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 1441; the nucleotide sequence of the full-length protein coding sequence of clone dw665_4 deposited under accession number ATCC 98818; or the nucleotide sequence of a mature protein coding sequence of clone dw665_4 deposited under accession number ATCC 98818. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological

activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:6, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 223 to amino acid 232 of SEQ ID NO:6.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:5, but excluding the poly(A) tail at the 3' end of SEQ ID NO:5; and

(ab) the nucleotide sequence of the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

(ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:5, but excluding the poly(A) tail at the 3' end of SEQ ID NO:5; and

(bb) the nucleotide sequence of the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:5, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:5 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:5 , but excluding the poly(A) tail at the 3' end of SEQ ID NO:5. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:5 from nucleotide 71 to nucleotide 1441, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:5 from nucleotide 71 to nucleotide 1441, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:5 from nucleotide 71 to nucleotide 1441. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 1441, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 1441, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 1441.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) a fragment of the amino acid sequence of SEQ ID NO:6, the fragment comprising eight contiguous amino acids of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:6, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 223 to amino acid 232 of SEQ ID NO:6.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial,

yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier. Preferably, the medical condition is selected from the group consisting of defects in cartilage, bone, or connective tissue formation, and damage to cartilage, bone, or connective tissue; more preferably, the medical condition is selected from the group consisting of broken bones; congenital, trauma-induced, or oncologic-resection-induced craniofacial defects; periodontal disease; defects in the periodontal ligament or attachment apparatus; damage to the periodontal ligament or attachment apparatus; osteoporosis; burns; incisions; and ulcers. Preferably, the protein of the present invention comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:4 from amino acid 35 to amino acid 1036, SEQ ID NO:4 from amino acid 637 to amino acid 1036, SEQ ID NO:6, SEQ ID NO:6 from amino acid 28 to amino acid 457, and SEQ ID NO:6 from amino acid 29 to amino acid 457.

Chordin-related proteins may be further characterized by the ability to demonstrate effects upon the growth and/or differentiation of embryonic cells and/or stem cells. Thus, the proteins or compositions of the present invention may also be useful for treating cell populations, such as embryonic cells or stem cell populations, to enhance or enrich the growth and/or differentiation of the cells.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

Figure 2 is a schematic representation of the chordin cysteine repeats (grey boxes) found in human Chordin protein, DW665_4 protein, and DJ167_19 protein (which has a partial chordin cysteine repeat at its amino teminus); the consensus amino acid sequence of the chordin cysteine repeat (SEQ ID NO:9) is also shown.

Figure 3 shows an agarose gel of reverse-transcriptase polymerase chain reaction (PCR) products indicating the relative levels of DJ167_19 mRNA expression in different tissues.

Figure 4 shows a northern blot of mRNA from different tissues probed with DJ167_19 sequences.

Figure 5 shows a dot blot of RNA from different tissues probed with DJ167_19 sequences; the DJ167_19 sequences appear to be most abundantly expressed in placenta (the dot at grid location F4).

Figure 6 shows an agarose gel of reverse-transcriptase polymerase chain reaction (PCR) products indicating the relative levels of DW665_4 mRNA expression in different tissues.

Figure 7 shows the results of experiments demonstrating expression of DW665_4 protein in COS cells and in CHO cells. Panel A shows a polyacrylamide gel of ^{35}-S-labeled proteins in conditioned medium from COS cells first transfected with either the pED vector (negative control), a DNA construct encoding human Chordin, or a DNA construct encoding DW665_4 protein. Panel B shows a western blot of DW665_4 proteins expressed in CHO cells.

Figure 8 is a schematic representation of the extent of binding between an N-terminal fragment of the DW665_4 protein and different members of the BMP protein family, as measured using a BIACORE instrument to detect changes in surface plasmon resistance.

Figure 9 shows the results of injecting DW665_19 RNA into the ventral blastomeres of 8-cell *Xenopus* embryos. Panel A is a mock-injected wild-type (WT) control. Panels B and C show the formation of primary (1°) and secondary (2°) axes in embryos injected with DW665_19 RNA, including the duplication of structures such as eyes (ey) and cement glands (cg).

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "dj167_2"

A polynucleotide of the present invention has been identified as clone "dj167_2". dj167_2 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dj167_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dj167_2 protein").

The nucleotide sequence of dj167_2 as presently determined is reported in SEQ ID NO:1, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dj167_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dj167_2 should be approximately 1550 bp.

The nucleotide sequence disclosed herein for dj167_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dj167_2 demonstrated at least some similarity with sequences identified as H49161 (yq18d05.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 274208 5'), L12350 (Human thrombospondin 2 (THBS2) mRNA, complete cds), T98917 (ye66b03.s1 Homo sapiens cDNA clone 122669 3' similar to SP:TSP1_CHICK P35440 THROMBOSPONDIN 1), and X87620 (B.taurus mRNA for complete thrombospondin). The predicted amino acid sequence disclosed herein for dj167_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dj167_2 protein demonstrated at least some similarity to sequences identified as L12350 (thrombospondin 2 [Homo sapiens]), M60853 (thrombospondin [Gallus gallus]), R40823 (Human thrombospondin 1), U48245 (protein kinase C-binding protein Nel [Rattus norvegicus]), X87620 (thrombospondin [Bos taurus]), and Z71178 (B0024.14 [Caenorhabditis elegans]). Based upon sequence similarity, dj167_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three potential transmembrane domains within the dj167_2 protein sequence, centered around amino acids 140, 215, and 315 of SEQ ID NO:2, respectively.

Clone "dj167_19"

A polynucleotide of the present invention has been identified as clone "dj167_19". dj167_19 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dj167_19 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dj167_19 protein").

The nucleotide sequence of dj167_19 as presently determined is reported in SEQ ID NO:3, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dj167_19 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 22 to 34 of SEQ ID NO:4 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 35. Due to the hydrophobic nature

of the predicted leader/signal sequence, it is likely to act as a transmembrane domain should the predicted leader/signal sequence not be separated from the remainder of the dj167_19 protein. The dj167_19 clone is related to that of dj167_2, and extends further 5'. The dj167_19 clone appears to contain coding sequences for chorionic somatomammotropin in the opposite orientation at its 5' end between Sfi restriction sites (at nucleotides 16 and 839 of SEQ ID NO:3). The dj167_2 and dj167_19 clones may represent alternatively spliced messenger RNA molecules encoding two different forms of a secreted protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dj167_19 should be approximately 4500 bp.

Analysis of the dj167_19 amino acid sequence (SEQ ID NO:4) reveals the following domains: IGFBP cysteine-rich domain at amino acids 60-75; VWF-B cysteine-rich domains at amino acids 174-210, 212-247, 255-291, and 293-328; Chordin cysteine-rich domains at amino acids 336-390, 403-456, 608-662, 679-734, 753-808, and 819-873; Antistasin (protease inhibitor) cysteine-rich domains at amino acids 469-498, 505-532, 539-564, and 567-592; RGD cell attachment sequence at amino acids 314-316, and Asn glycosylation sites at amino acids 71, 113, 330, 474, and 746. In addition, the amino acid sequence of SEQ ID NO:4 from amino acid 938 to amino acid 960 appears to be a transmembrane domain. The cysteine-rich domains listed above are similar to domains found in the C domain of Von Willebrand Factor (VWF), and in procollagen and thrombospondin. Antistasin, isolated from leeches, is a potent inhibitor of blood coagulation factor Xa.

DJ167_19 protein appears to be a unique membrane-bound BMP-binding protein. The dj167_19 transcript is expressed in a variety cell types, including kidney, pancreas, spleen, and ovary, and is most abundantly expressed in placental tissue. DJ167_19 protein may be an antagonist of BMP activity like Chordin; however, DJ167_19 protein is distinct from Chordin and the DW665_4 protein described below in that it possesses a heterogeneous protein domain structure, while the only protein domains identified so far in the Chordin and DW665_4 proteins are the chordin cysteine-rich repeats. Therefore, DJ167_19 protein may interact with other proteins in addition to BMP protein family members. The presence of antistasin domains in the DJ167_19 protein may indicate that DJ167_19 protein is regulated by proteolysis *in vivo*: the antistasins are "sacrificial" protease inhibitors. Like Chordin, DJ167_19 protein may therefore be a "reversible", proteolysis-regulated inhibitor of BMP activity.

Clone "dw665_4"

A polynucleotide of the present invention has been identified as clone "dw665_4". dw665_4 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dw665_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dw665_4 protein").

The nucleotide sequence of dw665_4 as presently determined is reported in SEQ ID NO:5, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dw665_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 15 to 27 of SEQ ID NO:6 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 28. Amino acids 16 to 28 of SEQ ID NO:6 are also a predicted leader/signal sequence, with the predicted mature amino acid sequence in that case beginning at amino acid 29. Due to the hydrophobic nature of these predicted leader/signal sequences, each is likely to act as a transmembrane domain should it not be separated from the remainder of the dw665_4 protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dw665_4 should be approximately 3750 bp.

The nucleotide sequence disclosed herein for dw665_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dw665_4 demonstrated at least some similarity with sequences identified as AA029053 (zk09f06.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 470051 3'), H77289 (EST27o17 WATM1 Homo sapiens cDNA clone 27o17, mRNA sequence), and T21722 (Human gene signature HUMGS03170). The predicted amino acid sequence disclosed herein for dw665_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dw665_4 protein demonstrated at least some similarity to sequences identified as L35764 (chordin [Xenopus laevis]) and W31559 (Xenopus frog protein "chordin"). Analysis of motifs within the predicted dw665_4 protein revealed the presence of Chordin cysteine-rich domains at amino acids 37-99, 115-178, and 260-322 of SEQ ID NO:6; an 'RGD" cell-attachment sequence (at amino acids 179-181 of SEQ ID NO:6), which in some

proteins has been shown to play a role in cell adhesion; and Asn glycosylation sites at amino acids 118 and 291. Based upon sequence similarity, dw665_4 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of dw665_4 indicates that it may contain an AC repetitive element.

DW665_4 protein is a novel human chordin-related protein. dw665_4 transcripts are expressed in many tissues including kidney, adrenal gland, and prostate tissues, and are most abundantly expressed in pancreas; however, little or no dw665_4 transcript expression is observed in liver or peripheral blood cells (see Figure 6). The sensitivity of CHO-prodced DW665_4 protein to proteolysis (see Figure 7) suggests that, like Chordin, DW665_4 protein may also be a "reversible", proteolysis-regulated inhibitor of BMP activity.

Deposit of Clones

Clones dj167_2 and dw665_4 were deposited on July 16, 1998 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original composite deposit under the Budapest Treaty and were given the accession number ATCC 98818, from which each clone comprising a particular polynucleotide is obtainable.

Clone dj167_19 was deposited on February 5, 1999 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original deposit under the Budapest Treaty and was given the accession number ATCC 207090, from which the dj167_19 clone comprising a particular polynucleotide is obtainable.

Each clone has been transfected into separate bacterial cells (*E. coli*) in these deposits. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Figures 1A and 1B, respectively. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped"

(i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of an oligonucleotide probe that was used to isolate or to sequence each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
dj167_2	SEQ ID NO:7
dw665_4	SEQ ID NO:8

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as, for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with γ-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated

by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as

immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

The chromosomal location corresponding to the polynucleotide sequences disclosed herein may also be determined, for example by hybridizing appropriately labeled polynucleotides of the present invention to chromosomes *in situ*. It may also be possible to determine the corresponding chromosomal location for a disclosed polynucleotide by identifying significantly similar nucleotide sequences in public databases, such as expressed sequence tags (ESTs), that have already been mapped to particular chromosomal locations. For at least some of the polynucleotide sequences disclosed herein, public database sequences having at least some similarity to the

polynucleotide of the present invention have been listed by database accession number. Searches using the GenBank accession numbers of these public database sequences can then be performed at an Internet site provided by the National Center for Biotechnology Information having the address http://www.ncbi.nlm.nih.gov/UniGene/, in order to identify "UniGene clusters" of overlapping sequences. Many of the "UniGene clusters" so identified will already have been mapped to particular chromosomal sites.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* 15(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* 62(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* 58: 1-39; all of which are incorporated by reference herein). The desired change in gene expression can also be achieved through the use of double-stranded ribonucleotide molecules having some complementarity to the mRNA transcribed from the gene, and which interfere with the transcription, stability, or expression of the mRNA ("RNA intereference" or "RNAi"; Fire *et al.*, 1998, *Nature* 391 (6669): 806-811; Montgomery *et al.*, 1998, *Proc. Natl. Acad. Sci. USA* 95 (26): 15502-15507; and Sharp, 1999, *Genes Dev.* 13 (2): 139-141; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* 14(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988,

Nature 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms, part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information. For example, the TopPredII computer program can be used to predict the location of transmembrane domains in an amino acid sequence, domains which are described by the location of the center of the transmsmbrane domain, with at least ten transmembrane amino acids on each side of the reported central residue(s).

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

In particular, sequence identity may be determined using WU-BLAST (Washington University BLAST) version 2.0 software, which builds upon WU-BLAST version 1.4, which in turn is based on the public domain NCBI-BLAST version 1.4 (Altschul and Gish, 1996, Local alignment statistics, Doolittle *ed.*, *Methods in Enzymology* **266**: 460-480; Altschul *et al.*, 1990, Basic local alignment search tool, *Journal of*

Molecular Biology **215**: 403-410; Gish and States, 1993, Identification of protein coding regions by database similarity search, *Nature Genetics* **3**: 266-272; Karlin and Altschul, 1993, Applications and statistics for multiple high-scoring segments in molecular sequences, *Proc. Natl. Acad. Sci. USA* **90**: 5873-5877; all of which are incorporated by reference herein). WU-BLAST version 2.0 executable programs for several UNIX platforms can be downloaded from ftp://blast.wustl.edu/blast/executables. The complete suite of search programs (BLASTP, BLASTN, BLASTX, TBLASTN, and TBLASTX) is provided at that site, in addition to several support programs. WU-BLAST 2.0 is copyrighted and may not be sold or redistributed in any form or manner without the express written consent of the author; but the posted executables may otherwise be freely used for commercial, nonprofit, or academic purposes. In all search programs in the suite -- BLASTP, BLASTN, BLASTX, TBLASTN and TBLASTX -- the gapped alignment routines are integral to the database search itself, and thus yield much better sensitivity and selectivity while producing the more easily interpreted output. Gapping can optionally be turned off in all of these programs, if desired. The default penalty (Q) for a gap of length one is Q=9 for proteins and BLASTP, and Q=10 for BLASTN, but may be changed to any integer value including zero, one through eight, nine, ten, eleven, twelve through twenty, twenty-one through fifty, fifty-one through one hundred, etc. The default per-residue penalty for extending a gap (R) is R=2 for proteins and BLASTP, and R=10 for BLASTN, but may be changed to any integer value including zero, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve through twenty, twenty-one through fifty, fifty-one through one hundred, etc. Any combination of values for Q and R can be used in order to align sequences so as to maximize overlap and identity while minimizing sequence gaps. The default amino acid comparison matrix is BLOSUM62, but other amino acid comparison matrices such as PAM can be utilized.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence

identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa,* and *Equus caballus,* for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* 22: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* 25: 682-690; Lyons *et al.*, 1997, *Nature Genetics* 15: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* 13(10): 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides that hybridize under reduced stringency conditions, more preferably stringent conditions, and most preferably highly

stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

Proteins of the invention also include other modified forms of the protein. It is known, for example, that numerous conservative amino acid substitutions are possible without significantly modifying the structure and conformation of a protein, thus maintaining the biological properties as well. For example, it is recognized that conservative amino acid substitutions may be made among amino acids with basic side chains, such as lysine (Lys or K), arginine (Arg or R) and histidine (His or H); amino acids with acidic side chains, such as aspartic acid (Asp or D) and glutamic acid (Glu or E); amino acids with uncharged polar side chains, such as asparagine (Asn or N), glutamine (Gln or Q), serine (Ser or S), threonine (Thr or T), and tyrosine (Tyr or Y); and amino acids with nonpolar side chains, such as alanine (Ala or A), glycine (Gly or G), valine (Val or V), leucine (Leu or L), isoleucine (Ile or I), proline (Pro or P), phenylalanine (Phe or F), methionine (Met or M), tryptophan (Trp or W) and cysteine (Cys or C). Thus, these modifications and deletions of the native chordin-related proteins may be employed as biologically active substitutes for naturally-occurring chordin-related proteins and other polypeptides in therapeutic processes. It can be readily determined whether a given

variant of a chordin-related protein maintains the biological activity of chordin by subjecting chordin, the naturally-occurring chordin-related protein, and the variant of the chordin-related protein to the assays described in the examples.

Other specific mutations of the sequences of chordin-related proteins described herein involve modifications of glycosylation sites. These modifications may involve O-linked or N-linked glycosylation sites. For instance, the absence of glycosylation or only partial glycosylation results from amino acid substitution or deletion at asparagine-linked glycosylation recognition sites. The asparagine-linked glycosylation recognition sites comprise tripeptide sequences which are specifically recognized by appropriate cellular glycosylation enzymes. These tripeptide sequences are either asparagine-X-threonine or asparagine-X-serine, where X is usually any amino acid. A variety of amino acid substitutions or deletions at one or both of the first or third amino acid positions of a glycosylation recognition site (and/or amino acid deletion at the second position) results in non-glycosylation at the modified tripeptide sequence. Additionally, bacterial expression of a chordin-related protein will also result in production of a non-glycosylated protein, even if the glycosylation sites are left unmodified.

The isolated polynucleotide endcoing the protein of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida,* or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium,* or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.,* Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLabs (Beverly, MA), Pharmacia (Piscataway, NJ) and

Invitrogen Corporation (Carlsbad, CA), respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from the Eastman Kodak Company (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The purified expressed protein is substantially free from other proteinaceous materials with which it is co-produced, as well as from other contaminants. The recovered purified protein is contemplated to have the ability to bind to BMPs and hence to exhibit effects on cartilage, bone, and/or other connective tissue formation activity. Thus, the proteins of the invention may be further characterized by the ability to demonstrate effects on cartilage, bone, and/or other connective tissue formation activity in bone and cartilage formation and other assays described below. Chordin-related proteins may be further characterized by the ability to demonstrate effects upon the growth and/or differentiation of embryonic cells and/or stem cells. Thus, the proteins or compositions of the present invention may also be characterized by their ability to enhance, enrich, or otherwise influence the growth and/or differentiation of the cells.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

EXAMPLE 1: Tissue-Specific Expression of DJ167_19 and DW665_4 Transcripts

A. Reverse-Transcription (RT) Polymerase Chain Reaction (PCR)

A 96-well panel of poly-A-containing RNA molecules from different human tissues (OriGene Technologies, Inc., Rockville, MD) was subjected to RT PCR under standard conditions and using sets of PCR primers based on either the dj167_19 or the dw665_4 polynucleotide sequences (SEQ ID NO:3 or SEQ ID NO:5, respectively). The PCR products were run on agarose gels and the results are shown in Fig. 3 for dj167_19 and in Fig. 6 for dw665_4. The dj167_19 sequences appear to be most highly expressed in kidney, placenta, and pancreas, although at least some expression was observed in most tissue types. The dw665_4 sequences appear to be most highly expressed in kidney, adrenal gland, and pancreas, although at least some expression was observed in most tissue types.

B. Northern Blot

A northern blot was prepared using RNA isolated from a number of tissue types, then probed under stringent conditions (such as those presented in the above table) with

a probe containing dj167_19 sequences. The result is shown in Figure 4. In this experiment, the most abundant transcript(s) are seen in placenta, kidney, spleen, and ovary. A major band is seen to migrate between the markers at 4.4 kb and 7.5 kb: this band may be estimated to be approximately 6 ± 1 kb in size. A faint band may be visible between the markers at 2.4 kb and 4.4 kb: this may represent a minor band estimated to be approximately 3.5 ± 0.5 kb in size, or it may simply be due to degradation of the major band.

C. Dot Blot

A dot blot was prepared using RNA isolated from a number of tissue types, then probed under stringent conditions (such as those presented in the above table) with a probe containing dj167_19 sequences; the results are shown in Figure 5. The grid locations of RNA dots reacting to different degrees to the DJ167_19 probe are as follows: row A: brain; dot B7: spinal cord; dot C2: aorta; dot C3: skeletal muscle; dot C4: colon; dot C6: uterus; dot C7: prostate; dot C8: stomach; dot D2: ovary; dot D5: adrenal gland; dot D6: thyroid; dot D7: salivary gland; dot D8: mammary gland; dot E1: kidney; dot E3: small intestine; dot E4: spleen; dot E5: thymus; dot E6: peripheral blood lymphocytes (PBL); dot E7: lymph node; dot F1: appendix; dot F2: lung; dot F3: trachea; dot F4: placenta; dot G2: fetal heart; dot G3: fetal kidney; dot G5: fetal spleen; dot G7: fetal lung; unlabeled row below row G: negative controls. In this experiment, the dj167_19 sequences appear to be expressed most abundantly in placenta.

EXAMPLE 2: Expression and purification of DW665_4 Protein

Figure 7 shows the results of experiments demonstrating expression of DW665_4 protein in COS cells and in CHO cells. Panel A shows a polyacrylamide gel of conditioned medium from COS cells first transfected with either the pED vector (negative control), a DNA construct encoding human Chordin, or a DNA construct encoding DW665_4 protein, and then labeled with ^{35}-S. A single DW665_4 protein band is observed in this experiment.

Panel B of Figure 7 shows a western blot of His-tagged DW665_4 proteins expressed in CHO cells; the proteins were detected using the PentaHis anti-His antibody as primary antibody and mouse HRP (horse radish peroxidase) antibody as the secondary antibody. This experiment indicates that DW665_4 proteins are cleaved in CHO cells to

form two smaller (approximately between 25 kDa and 30 kDa) fragments: aprotinin did not inhibit this proteolysis in CHO cells, but culturing the cells with a pellet of protease inhibitors allowed some of the larger DW665_4 His-tagged protein to be produced (indicated as "Full Length" in Figure 7). The larger His-tagged DW665_4 protein ("HEKDW665") and the N-terminal fragment ("DW665-N") were retained on a Ni-NTA column and eluted at approximately the same time. The C-terminal fragment ("DW665-C") was eluted separately from the Ni-NTA column. Further purification of the HEKDW665 and DW665-N proteins was performed on a MonoS column: only DW665-N was found in the eluate, and stripping the column did not elute the larger HEKDW665 protein. N-terminal sequencing of the C-terminal fragment DW665-C indicated that it has N-termini at the Serine residue at position 230 of SEQ ID NO:6 and at the Glycine residue at position 233 of SEQ ID NO:6; the C-terminal fragment DW665-C therefore has one of the three Chordin cysteine repeats. Based on this data, the N-terminal fragment DW665-N is expected to have a C-terminus extending no further than to an amino acid between position 229 and 232 of SEQ ID NO:6, and the N-terminal fragment DW665-N is expected to retain two of the three Chordin cysteine repeats.

EXAMPLE 3: Binding to BMP Family Members and BMP Inhibition Activity

Figure 8 is a schematic representation of the extent of binding between an N-terminal fragment of the DW665_4 protein (DW665-N) and different members of the BMP protein family, as measured using a BIACORE instrument to detect changes in surface plasmon resistance. The BIACORE binding experiments indicate that the N-terminal fragment of the DW665_4 protein has a Chordin-like protein-binding profile, and binds to BMP-2, BMP-4, BMP-7, and GDF-5, and to a lesser degree to BMP-12 and BMP-13. However, this N-terminal fragment of DW665_4 does not seem to inhibit BMP-2 in the W20 bioassay (see below). This result suggests that the third Chordin cysteine repeat present in the DW665-C fragment may be necessary for BMP-inhibitory activity.

EXAMPLE 4: DW665_4 Induces Axis Duplication in *Xenopus* Embryos

DW665_4 RNA (250 pg of RNA per injection) was injected into the ventral blastomeres of 8-cell *Xenopus* embryos; the results are shown in Figure 9. Panel A is a mock-injected wild-type (WT) control. Panels B and C show the formation of primary

(1°) and secondary (2°) axes in embryos injected with DW665_19 RNA, including the duplication of structures such as eyes (ey) and cement glands (cg). Of the embryos injected with DW665_4 RNA (n = 40), 60% exhibited secondary axis formation, and some duplications were fairly complete, including eyes and cement glands as shown in Panel C of Figure 9. Chordin RNA produces a similar result when injected into *Xenopus* embryos.

EXAMPLE 5: Bioassays

A. W-20 BIOASSAYS

1. Description of W-20 cells

Use of the W-20 bone marrow stromal cells as an indicator cell line is based upon the conversion of these cells to osteoblast-like cells after treatment with a BMP protein [Thies et al, Journal of Bone and Mineral Research, 5:305 (1990); and Thies et al, Endocrinology, 130:1318 (1992)]. Specifically, W-20 cells are a clonal bone marrow stromal cell line derived from adult mice by researchers in the laboratory of Dr. D. Nathan, Children's Hospital, Boston, MA. Treatment of W-20 cells with certain BMP proteins results in (1) increased alkaline phosphatase production, (2) induction of PTH stimulated cAMP, and (3) induction of osteocalcin synthesis by the cells. While (1) and (2) represent characteristics associated with the osteoblast phenotype, the ability to synthesize osteocalcin is a phenotypic property only displayed by mature osteoblasts. Furthermore, to date we have observed conversion of W-20 stromal cells to osteoblast-like cells only upon treatment with BMPs. In this manner, the *in vitro* activities displayed by BMP treated W-20 cells correlate with the *in vivo* bone forming activity known for BMPs.

Below two *in vitro* assays useful in comparison of BMP activities of novel osteoinductive molecules are described.

2. W-20 Alkaline Phosphatase Assay Protocol

W-20 cells are plated into 96 well tissue culture plates at a density of 10,000 cells per well in 200 µl of media (DME with 10% heat inactivated fetal calf serum, 2 mM glutamine and 100 Units/ml penicillin + 100 µg/ml streptomycin. The cells are allowed to attach overnight in a 95% air, 5% CO_2 incubator at 37°C. The 200 µl of media is removed from each well with a multichannel pipettor and replaced with an equal volume of test sample delivered in DME with 10% heat inactivated fetal calf serum, 2 mM glutamine and 1% penicillin-streptomycin. Test substances are assayed in triplicate. The

test samples and standards are allowed a 24 hour incubation period with the W-20 indicator cells. After the 24 hours, plates are removed from the 37°C incubator and the test media are removed from the cells. The W-20 cell layers are washed 3 times with 200 µl per well of calcium/magnesium free phosphate buffered saline and these washes are discarded. 50 µl of glass distilled water is added to each well and the assay plates are then placed on a dry ice/ethanol bath for quick freezing. Once frozen, the assay plates are removed from the dry ice/ethanol bath and thawed at 37°C. This step is repeated 2 more times for a total of 3 freeze-thaw procedures. Once complete, the membrane bound alkaline phosphatase is available for measurement. 50 µl of assay mix (50 mM glycine, 0.05% Triton X-100, 4 mM $MgCl_2$, 5 mM p-nitrophenol phosphate, pH = 10.3) is added to each assay well and the assay plates are then incubated for 30 minutes at 37°C in a shaking waterbath at 60 oscillations per minute. At the end of the 30 minute incubation, the reaction is stopped by adding 100 µl of 0.2 N NaOH to each well and placing the assay plates on ice. The spectrophotometric absorbance for each well is read at a wavelength of 405 nanometers. These values are then compared to known standards to give an estimate of the alkaline phosphatase activity in each sample. For example, using known amounts of p-nitrophenol phosphate, absorbance values are generated. This is shown in Table I.

Table I

Absorbance Values for Known Standards
of P-Nitrophenol Phosphate

P-nitrophenol phosphate umoles	Mean absorbance (405 nm)
0.000	0
0.006	0.261 +/- .024
0.012	0.521 +/- .031
0.018	0.797 +/- .063
0.024	1.074 +/- .061
0.030	1.305 +/- .083

Absorbance values for known amounts of BMPs can be determined and converted to µmoles of p-nitrophenol phosphate cleaved per unit time as shown in Table II.

Table II

Alkaline Phosphatase Values for W-20 Cells Treating with BMP-2

BMP-2 concentration ng/ml	Absorbance Reading 405 nmeters	umoles substrate per hour
0	0.645	0.024
1.56	0.696	0.026
3.12	0.765	0.029
6.25	0.923	0.036
12.50	1.121	0.044
25.0	1.457	0.058
50.0	1.662	0.067
100.0	1.977	0.080

These values are then used to compare the activities of known amounts of BMP-16 to BMP-2.

3. Osteocalcin RIA Protocol

W-20 cells are plated at 10^6 cells per well in 24 well multiwell tissue culture dishes in 2 mls of DME containing 10% heat inactivated fetal calf serum, 2 mM glutamine. The cells are allowed to attach overnight in an atmosphere of 95% air 5% CO_2 at 37°C. The next day the medium is changed to DME containing 10% fetal calf serum, 2 mM glutamine and the test substance in a total volume of 2 ml. Each test substance is administered to triplicate wells. The test substances are incubated with the W-20 cells for a total of 96 hours with replacement at 48 hours by the same test medias. At the end of 96 hours, 50 µl of the test media is removed from each well and assayed for osteocalcin production using a radioimmunoassay for mouse osteocalcin. The details of the assay are described in the kit manufactured by Biomedical Technologies Inc., 378 Page Street, Stoughton, MA 02072. Reagents for the assay are found as product numbers BT-431 (mouse osteocalcin standard), BT-432 (Goat anti-mouse Osteocalcin), BT-431R (iodinated mouse osteocalcin), BT-415 (normal goat serum) and BT-414 (donkey anti goat IgG). The RIA for osteocalcin synthesized by W-20 cells in response to BMP treatment is carried out as described in the protocol provided by the manufacturer.

The values obtained for the test samples are compared to values for known standards of mouse osteocalcin and to the amount of osteocalcin produced by W-20 cells in response to challenge with known amounts of BMP-2. The values for BMP-2 induced osteocalcin synthesis by W-20 cells is shown in Table III.

Table III

Osteocalcin Synthesis by W-20 Cells

BMP-2 Concentration ng/ml	Osteocalcin Synthesis ng/well
0	0.8
2	0.9
4	0.8
8	2.2
16	2.7
31	3.2
62	5.1
125	6.5
250	8.2
500	9.4
1000	10.0

B. ROSEN MODIFIED SAMPATH-REDDI ASSAY

A modified version of the rat bone formation assay described in Sampath and Reddi, Proc. Natl. Acad. Sci. USA, 80:6591-6595 (1983) is used to evaluate bone and/or cartilage and/or other connective tissue activity of BMP proteins. This modified assay is herein called the Rosen-modified Sampath-Reddi assay. The ethanol precipitation step of the Sampath-Reddi procedure is replaced by dialyzing (if the composition is a solution) or diafiltering (if the composition is a suspension) the fraction to be assayed against water. The solution or suspension is then equilibrated to 0.1% TFA. The resulting solution is added to 20 mg of rat matrix. A mock rat matrix sample not treated with the protein serves as a control. This material is frozen and lyophilized and the resulting powder enclosed in #5 gelatin capsules. The capsules are implanted subcutaneously in the abdominal thoracic area of 21-49 day old male Long Evans rats. The implants are removed after 7-14 days. Half of each implant is used for alkaline phosphatase analysis [see, Reddi et al, Proc. Natl. Acad. Sci., 69:1601 (1972)].

The other half of each implant is fixed and processed for histological analysis. 1 μm glycolmethacrylate sections are stained with Von Kossa and acid fuschin to score the amount of induced bone and cartilage and other connective tissue formation present in each implant. The terms +1 through +5 represent the area of each histological section of an implant occupied by new bone and/or cartilage cells and matrix. A score of +5 indicates that greater than 50% of the implant is new bone and/or cartilage produced as a direct result of protein in the implant. A score of +4, +3, +2, and +1 would indicate that greater than 40%, 30%, 20% and 10% respectively of the implant contains new cartilage and/or bone.

Alternatively, the implants are inspected for the appearance of tissue resembling embryonic tendon, which is easily recognized by the presence of dense bundles of fibroblasts oriented in the same plane and packed tightly together. [Tendon/ligament-like tissue is described, for example, in Ham and Cormack, Histology (JB Lippincott Co. (1979), pp. 367-369, the disclosure of which is hereby incorporated by reference]. These findings may be reproduced in additional assays in which tendon/ligament-like tissues are observed in the chordin-related protein containing implants. The chordin-related proteins of this invention may be assessed for activity on this assay.

C. Embryonic Stem Cell Assay

In order to assay the effects of the chordin-related proteins of the present invention, it is possible to assay the growth and differentiation effects *in vitro* on a number of available embryonic stem cell lines. One such cell line is ES-E14TG2, which is available from the American Type Culture Collection in Rockville, Md.

In order to conduct the assay, cells may be propagated in the presence of 100 units of LIF to keep them in an undifferentiated state. Assays are setup by first removing the LIF and aggregating the cells in suspension, in what is known as embryoid bodies. After 3 days the embryoid bodies are plated on gelatin coated plates (12 well plates for PCR analysis, 24 well plates for immunocytochemistry) and treated with the proteins to be assayed. Cells are supplied with nutrients and treated with the protein factor every 2-3 days. Cells may be adapted so that assays may be conducted in media supplemented with 15% Fetal Bovine Serum (FBS) or with CDM defined media containing much lower amounts of FBS.

At the end of the treatment period (ranging from 7-21 days) RNA is harvested from the cells and analyzed by quantitative multiplex PCR for the following markers:

Brachyury, a mesodermal marker, *AP-2*, an ectodermal marker, and *HNF-3α* an endodermal marker. Through immunocytochemistry, it is also possible to detect the differentiation of neuronal cells (glia and neurons), muscle cells (cardiomyocytes, skeletal and smooth muscle), and various other phenotype markers such as proteoglycan core protein (cartilage), and cytokeratins (epidermis). Since these cells have a tendency to differentiate autonomously when LIF is removed, the results are always quantitated by comparison to an untreated control.

EXAMPLE 6: BMP Binding Assays

The chordin-related polypeptides of the present invention may be assayed for binding to BMPs, other TGF-β proteins, or other ligands in any manner known in the art, including the following methods:

Ligand Blotting: The binding protein [chordin-related polypeptide] is run on SDS-PAGE, transferred to a membrane (such as a Western blot) and probed with iodinated ligand. Fukui et al., Developmental Biology, 159:131-139 (1993).

Gel Filtration: The binding protein [chordin-related polypeptide] is incubated with iodinated ligand and and ligand-binding protein complex is separated from unbound species by size using gel filtration. Vaughn and Vale, Endocrinology, 132:2038-2050 (1993).

Cross-Linking: The binding protein [chordin-related polypeptide] is incubated with iodinated ligand and covalently coupled with chemical cross-linker. The reaction mix is run on SDS-PAGE. Autoradiography will reveal complex formation via binding of ligand to binding protein. Vaughn and Vale, Endocrinology, 132:2038-2050 (1993).

Immunoprecipitation: The binding protein [chordin-related polypeptide] is incubated with iodinated ligand and covalently coupled with chemical cross-linker. The reaction mix is then immunoprecipitated with ligand antibody. The immunoprecipitate is run on SDS-PAGE. Vaughn and Vale, Endocrinology, 132:2038-2050 (1993).

Gel Shift: The binding protein [chordin-related polypeptide] is incubated with iodinated ligand and run on non-denaturing agarose gel. The complex is identified by autoradiography. Krumment at al., Endocrinology, 132:431-443 (1993).

Radioreceptor Binding Assay: The ligand is iodinated and specific activity is determined. The cell surface receptor binding assay described in Massague, Methods in Enzymology, 46:174-195 (1987) is performed using 10T1/2 cells, or other suitable cell line. The cells are

allowed to reach confluency in suitable medium, rinsed, and incubated with iodinated ligand containing increasing concentrations of binding protein [chordin-related polypeptide] at room temperature for one hour. The plates are chilled and rinsed. The bound iodinated ligand is solubilized with solubilization buffer and counted with a gamma counter. Massague, *Id.*

The above references are hereby incorporated herein by reference for their full disclosure of the methods and materials useful in the above procedures.

USES AND BIOLOGICAL ACTIVITIES: TISSUE INDUCTION AND REPAIR

A protein of the present invention, which induces or influences cartilage and/or bone and/or other connective tissue formation, may have application in the healing of bone fractures and cartilage or other connective tissue defects in humans and other animals. Such a preparation employing a chordin-related protein may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery. A chordin-related protein may be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, affect or stimulate growth or differentiation of bone-forming cells and their progenitor cells or induce differentiation of progenitors of bone-forming cells, and may also support the regeneration of the periodontal ligament and attachment apparatus, which connects bone and teeth. Chordin-related polypeptides of the invention may also be useful in the treatment of systemic conditions such as osteoporosis, and under certain circumstances, to augment or inhibit the effects of osteogenic, cartilage-inducing, and bone inducing factors. In addition to the TGF-β superfamily of proteins, a variety of osteogenic, cartilage-inducing and bone-inducing factors have been described. See, e.g., European patent applications 148,155 and 169,016 for discussions thereof.

The proteins of the invention may also be used in wound healing and related tissue repair. The types of wounds include, but are not limited to burns, incisions and ulcers. (See, e.g. PCT Publication WO84/01106 for discussion of wound healing and related tissue repair). It is further contemplated that proteins of the invention may affect neuronal, astrocytic, and glial cell survival and therefore be useful in transplantation and treatment of conditions exhibiting a decrease in neuronal survival and repair. The

proteins of the invention may further be useful for the treatment of conditions related to other types of tissue, such as nerve, epidermis, muscle, and other organs such as liver, brain, lung, cardiac, pancreas, and kidney tissue. The proteins of the present invention may further be useful for the treatment of relatively undifferentiated cell populations, such as embryonic cells, or stem cells, to enhance growth and/or differentiation of the cells; such enhancement of growth and/or differentiation of these cells may particularly be carried out on isogenic or allogenic cells *ex vivo*, with subsequent reintroduction of the treated cells to the patient. The proteins of the present invention may also have value as a dietary supplement, or as a component of cell culture media. For this use, the proteins may be used in intact form, or may be predigested to provide a more readily absorbed supplement.

The proteins of the invention may also have other useful properties characteristic of the TGF-β superfamily of proteins. Such properties include angiogenic, chemotactic, and/or chemoattractant properties, and effects on cells including induction or inhibition of collagen synthesis, fibrosis, differentiation responses, cell proliferative responses, and responses involving cell adhesion, migration, and extracellular matrices. These properties make the proteins of the invention potential agents for wound healing, reduction of fibrosis, and reduction of scar tissue formation.

Chordin-related proteins, alone or complexed with monomers, homodimers, or heterodimers of BMPs, with members of the TGF-β superfamily of proteins, or with inhibin-α proteins or inhibin-β proteins, the chordin-related protein is expected to demonstrate effects on the production of follicle stimulating hormone (FSH), as described further herein. It is recognized that FSH stimulates the development of ova in mammalian ovaries (Ross et al., in Textbook of Endocrinology, ed. Williams, p. 355 (1981) and that excessive stimulation of the ovaries with FSH will lead to multiple ovulations. FSH is also important in testicular function. Thus, chordin-related proteins may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in mammals. Chordin-related proteins may also be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. Chordin-related proteins may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs. It is further

contemplated that chordin-related proteins may be useful in modulating hematopoiesis by inducing the differentiation of erythroid cells [see, e.g., Broxmeyer et al, *Proc. Natl. Acad. Sci. USA*, 85:9052-9056 (1988) or Eto et al, *Biochem. Biophys. Res. Comm.*, 142:1095-1103 (1987)], for suppressing the development of gonadal tumors [see, e.g., Matzuk et al., *Nature*, 360:313-319 (1992)] or for augmenting the activity of bone morphogenetic proteins [see, e.g., Ogawa et al., *J. Biol. Chem.*, 267:14233-14237 (1992)].

Chordin-related proteins proteins may be further characterized by their ability to modulate the release of follicle stimulating hormone (FSH) in established *in vitro* bioassays using rat anterior pituitary cells as described [see, e.g., Vale et al, *Endocrinology*, 91:562-572 (1972); Ling et al., *Nature*, 321:779-782 (1986) or Vale et al., *Nature*, 321:776-779 (1986)]. It is contemplated that the chordin-related protein of the invention may bind to TGF-β proteins, which will have different effects depending upon whether they are in homodimeric or heterodimeric form. TGF-β proteins when found as a heterodimer with inhibin α or inhibin β chains, will exhibit regulatory effects, either stimulatory or inhibitory, on the release of follicle stimulating hormone (FSH), from anterior pituitary cells as described [Ling et al., *Nature*, 321:779-782 (1986) or Vale et al., *Nature*, 321:776-779 (1986); Vale et al, *Endocrinology*, 91:562-572 (1972). Therefore, depending on the particular composition, it is expected that the chordin-related protein of the invention may have contrasting and opposite effects on the release of follicle stimulating hormone (FSH) from the anterior pituitary.

Activin A (the homodimeric composition of inhibin β_A) has been shown to have erythropoietic-stimulating activity [see e.g. Eto et al., *Biochem. Biophys. Res. Commun.*, 142:1095-1103 (1987) and Murata et al., *Proc. Natl. Acad. Sci. U.S.A.*, 85:2434-2438 (1988) and Yu et al., *Nature*, 330:765-767 (1987)]. It is contemplated that the chordin-related protein of the invention may have a similar erythropoietic-stimulating activity. This activity of the chordin-related protein may be further characterized by the ability of the chordin-related protein to demonstrate erythropoietin activity in the biological assay performed using the human K-562 cell line as described by [Lozzio et al., *Blood*, 45:321-334 (1975) and U.S. Pat. No. 5,071,834].

A further aspect of the invention is a therapeutic method and composition for repairing fractures and other conditions related to cartilage and/or bone and/or other connective tissue defects or periodontal diseases. The invention further comprises therapeutic methods and compositions for wound healing and tissue repair. Such compositions comprise a therapeutically effective amount of at least one of the chordin-

related proteins of the invention in a mixture with a pharmaceutically acceptable vehicle, carrier, or matrix. It is further contemplated that compositions of the invention may increase neuronal survival and therefore be useful in transplantation and treatment of conditions exhibiting a decrease in neuronal survival. Compositions of the invention may further include at least one other therapeutically useful agent, such as members of the TGF-β superfamily of proteins, which includes the BMP proteins BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6 and BMP-7, disclosed for instance in United States Patents 5,108,922; 5,013,649; 5,116,738; 5,106,748; 5,187,076; and 5,141,905; BMP-8, disclosed in PCT publication WO91/18098; BMP-9, disclosed in PCT publication WO93/00432; BMP-10, disclosed in PCT application WO94/26893; BMP-11, disclosed in PCT application WO94/26892, BMP-12 or BMP-13, disclosed in PCT application WO 95/16035, or BMP-15, disclosed in co-pending patent application, serial no. 08/446,924, filed on May 18, 1995; or BMP-16, disclosed in co-pending patent application, serial no. 715,202, filed on September 18, 1996. Other compositions which may also be useful include Vgr-2, and any of the growth and differentiation factors [GDFs], including those described in PCT applications WO94/15965; WO94/15949; WO95/01801; WO95/01802; WO94/21681; WO94/15966; WO95/10539; WO96/01845; WO96/02559 and others. Also useful in the present invention may be BIP, disclosed in WO94/01557; HP00269, disclosed in JP Publication number: 7-250688; and MP52, disclosed in PCT application WO93/16099. The disclosures of the above applications are hereby incorporated by reference herein.

Further, tyrosine kinase receptor genes and/or proteins, and/or soluble truncated versions thereof, may also be useful in compositions of the present invention, including the following receptors, or soluble truncated versions comprising the extracellular binding domains thereof: LTK, Toyoshima *et al.*, PNAS USA *90*:5404 (1993); TIE, Partanen *et al.*, Mol. Cell Biol *12*:1698 (1992); DTK, Crosier *et al.*, Growth Factors *11*:137 (1994); MER, Graham *et al.*, Cell Growth and Differentiation *5*:647 (1994); ALK, Morris *et al.*, Science *263*:1281 (1994); RYK, Tamagnone *et al.*, Oncogene *8*:2009 (1993); Paul *et al.*, Int. J Cell Cloning *10*:309 (1992); ROR1 and ROR2, Masiakowski and Carroll, J. Biol. Chem. *267*:26181 (1992); MuSK/Mlk/Nsk2, Valenzuela *et al.*, Neuron *15*:573 (1995); Ganju *et al.*, Oncogene *11*:281 (1995); TKT, Karn *et al.*, Oncogene *8*:3443 (1993); and DDR, Johnson *et al.*, PNAS USA *90*:5677 (1993). The disclosure of the above references is hereby incorporated by reference as if reproduced fully herein.

It is expected that human chordin-related proteins may exist in nature as homodimers or heterodimers. To promote the formation of dimers of chordin-related

proteins with increased stability, one can genetically engineer the DNA sequences encoding these proteins to provide one or more additional cysteine residues to increase potential dimer formation. The resulting DNA sequence would be capable of producing a "cysteine added variant" of the chordin-related protein. In a preferred embodiment, one would engineer the DNA sequences encoding the chordin-related protein so that one or more codons may be altered to a nucleotide triplet encoding a cysteine residue, such as TGT or TGC. Alternatively, one can produce "cysteine added variants" of chordin-related proteins by altering the sequence of the protein at the amino acid level by altering one or more amino acid residues to Cys. Production of "cysteine added variants" of proteins is described in United States Patent 5,166,322, the disclosure of which is hereby incorporated by reference.

It is expected that the proteins of the invention may act in concert with or perhaps synergistically with other related proteins and growth factors. Further therapeutic methods and compositions of the invention therefore comprise a therapeutic amount of at least one chordin-related protein of the invention with a therapeutic amount of at least one protein growth and/or differentiation factor, such as a member of the TGF-β superfamily of proteins, such as the BMP proteins disclosed in the applications described above. Such combinations may comprise chordin-related protein with separate molecules of the BMP proteins or heteromolecules comprised of different BMP moieties. For example, a method and composition of the invention may comprise a disulfide-linked dimer comprising a chordin-related protein subunit and a subunit from one of the "BMP" proteins described above. Thus, the present invention includes a purified chordin-related polypeptide which is a heterodimer wherein one subunit comprises the amino acid sequence of a chordin-related protein of the invention or a fragment thereof, and one subunit comprises an amino acid sequence for a bone morphogenetic protein selected from the group consisting of BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8, BMP-9, BMP-10, BMP-11, BMP-12 or BMP-13, disclosed in PCT application WO 95/16035, VGR-2, MP-52, BIP, the GDFs, HP-269, or BMP-15, disclosed in co-pending patent application, serial number 08/446,924, filed on May 18, 1995; or BMP-16, disclosed in co-pending patent application, serial no. 715,202, filed on September 18, 1996. A further embodiment may comprise a heterodimer of chordin-related moieties, for example of one of the human chordin-related proteins described herein and the *xenopus chordin* protein, which is the homologue of human chordin. Further, chordin-related proteins may be combined with other agents beneficial to the treatment of the bone and/or cartilage

and/or other connective tissue defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), fibroblast growth factor (FGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), activins, inhibins, and k-fibroblast growth factor (kFGF), parathyroid hormone (PTH), parathyroid hormone related peptide (PTHrP), leukemia inhibitory factor (LIB/HILA/DA), insulin-like growth factors (IGF-I and IGF-II). Portions of these agents may also be used in compositions of the present invention. The preparation and formulation of such physiologically acceptable protein compositions, having due regard to pH, isotonicity, stability and the like, is within the skill of the art. The therapeutic compositions are also presently valuable for veterinary applications due to the lack of species specificity in growth and differentiation factors such as chordin-related proteins. Particularly domestic animals and thoroughbred horses in addition to humans are desired patients for such treatment with the chordin-related proteins of the present invention.

ADDITIONAL USES AND BIOLOGICAL ACTIVITIES

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting

and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those described in Gyuris *et al.*, 1993, *Cell* 75: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein

or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious

diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to regulate immune responses in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated

through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce

antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or

in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J.

Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood

84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et

al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce

differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present

invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle

stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population

of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of

such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting

from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue

in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via antibody-dependent cell-mediated cytotoxicity (ADCC)). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11,

IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the

art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the

present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and

the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. As used herein, the term "antibody" includes without limitation a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a single-chain antibody, a CDR-grafted antibody, a humanized antibody, or fragments thereof which bind to the indicated protein. Such term also includes any other species derived from an antibody or antibody sequence which is capable of binding the indicated protein.

Antibodies to a particular protein can be produced by methods well known to those skilled in the art. For example, monoclonal antibodies can be produced by generation of antibody-producing hybridomas in accordance with known methods (see for example, Goding, 1983, Monoclonal antibodies: principles and practice, Academic Press Inc., New York; and Yokoyama, 1992, "Production of Monoclonal Antibodies" in Current Protocols in Immunology, Unit 2.5, Greene Publishing Assoc. and John Wiley & Sons). Polyclonal sera and antibodies can be produced by inoculation of a mammalian subject with the relevant protein or fragments thereof in accordance with known methods. Fragments of antibodies, receptors, or other reactive peptides can be produced from the corresponding antibodies by cleavage of and collection of the desired fragments in accordance with known methods (see for example, Goding, supra; and Andrew et al., 1992, "Fragmentation of Immunoglobulins" in Current Protocols in Immunology, Unit 2.8, Greene Publishing Assoc. and John Wiley & Sons). Chimeric antibodies and single chain antibodies can also be produced in accordance with known recombinant methods (see for example, 5,169,939, 5,194,594, and 5,576,184). Humanized antibodies can also be made from corresponding murine antibodies in accordance with well known methods (see for example, U.S. Patent Nos. 5,530,101, 5,585,089, and 5,693,762). Additionally, human antibodies may be produced in non-human animals such as mice that have been genetically altered to express human antibody molecules (see for example Fishwild *et al.*, 1996, *Nature Biotechnology* **14**: 845-851; Mendez *et al.*, 1997, *Nature Genetics* **15**: 146-156 (erratum *Nature Genetics*

16: 410); and U.S. Patents 5,877,397 and 5,625,126). Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987).

Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium

sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence of SEQ ID NO:1;

(b) the nucleotide sequence of SEQ ID NO:1 from nucleotide 157 to nucleotide 1356;

(c) the nucleotide sequence of the full-length protein coding sequence of clone dj167_2 deposited under accession number ATCC 98818;

(d) a nucleotide sequence encoding the full-length protein encoded by the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

(e) a nucleotide sequence encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(f) a nucleotide sequence encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2, the fragment comprising eight contiguous amino acids of SEQ ID NO:2;

(g) the nucleotide sequence of a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified by (a)-(d); and

(h) the nucleotide sequence of a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified by (a)-(d), and that has a length that is at least 25% of the length of SEQ ID NO:1.

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of a host cell in a suitable culture medium, wherein the host cell has been transformed with the polynucleotide of claim 2; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. An isolated polynucleotide encoding the protein of claim 6.

8. The polynucleotide of claim 7, wherein the polynucleotide comprises the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818.

9. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) a fragment of the amino acid sequence of SEQ ID NO:2, the fragment comprising eight contiguous amino acids of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone dj167_2 deposited under accession number ATCC 98818;

the protein being substantially free from other mammalian proteins.

10. The protein of claim 9, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

11. A composition comprising the protein of claim 9 and a pharmaceutically acceptable carrier.

12. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence of SEQ ID NO:3;

(b) the nucleotide sequence of SEQ ID NO:3 from nucleotide 1383 to nucleotide 4490;

(c) the nucleotide sequence of SEQ ID NO:3 from nucleotide 1485 to nucleotide 4490;

(d) the nucleotide sequence of SEQ ID NO:3 from nucleotide 3645 to nucleotide 4343;

(e) the nucleotide sequence of the full-length protein coding sequence of clone dj167_19 deposited under accession number ATCC 207090;

(f) a nucleotide sequence encoding the full-length protein encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

(g) the nucleotide sequence of a mature protein coding sequence of clone dj167_19 deposited under accession number ATCC 207090;

(h) a nucleotide sequence encoding a mature protein encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

(i) a nucleotide sequence encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(j) a nucleotide sequence encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4, the fragment comprising eight contiguous amino acids of SEQ ID NO:4;

(k) the nucleotide sequence of a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified by (a)-(h); and

(l) the nucleotide sequence of a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified by (a)-(h), and that has a length that is at least 25% of the length of SEQ ID NO:3.

13. The polynucleotide of claim 12 wherein said polynucleotide is operably linked to at least one expression control sequence.

14. A host cell transformed with the polynucleotide of claim 13.

15. The host cell of claim 14, wherein said cell is a mammalian cell.

16. A process for producing a protein encoded by the polynucleotide of claim 13, which process comprises:

(a) growing a culture of a host cell in a suitable culture medium, wherein the host cell has been transformed with the polynucleotide of claim 13; and

(b) purifying said protein from the culture.

17. A protein produced according to the process of claim 16.

18. An isolated polynucleotide encoding the protein of claim 17.

19. The polynucleotide of claim 18, wherein the polynucleotide comprises the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090.

20. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 637 to amino acid 1036;

(c) a fragment of the amino acid sequence of SEQ ID NO:4, the fragment comprising eight contiguous amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone dj167_19 deposited under accession number ATCC 207090;

the protein being substantially free from other mammalian proteins.

21. The protein of claim 10, wherein said protein comprises the amino acid sequence of SEQ ID NO:4.

22. A composition comprising the protein of claim 20 and a pharmaceutically acceptable carrier.

23. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence of SEQ ID NO:5;

(b) the nucleotide sequence of SEQ ID NO:5 from nucleotide 71 to nucleotide 1441;

(c) the nucleotide sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 1441;

(d) the nucleotide sequence of the full-length protein coding sequence of clone dw665_4 deposited under accession number ATCC 98818;

(e) a nucleotide sequence encoding the full-length protein encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

(f) the nucleotide sequence of a mature protein coding sequence of clone dw665_4 deposited under accession number ATCC 98818;

(g) a nucleotide sequence encoding a mature protein encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

(h) a nucleotide sequence encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a nucleotide sequence encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6, the fragment comprising eight contiguous amino acids of SEQ ID NO:6;

(j) the nucleotide sequence of a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified by (a)-(g); and

(k) the nucleotide sequence of a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified by (a)-(g), and that has a length that is at least 25% of the length of SEQ ID NO:5.

24. The polynucleotide of claim 23 wherein said polynucleotide is operably linked to at least one expression control sequence.

25. A host cell transformed with the polynucleotide of claim 24.

26. The host cell of claim 25, wherein said cell is a mammalian cell.

27. A process for producing a protein encoded by the polynucleotide of claim 24, which process comprises:

(a) growing a culture of a host cell in a suitable culture medium, wherein the host cell has been transformed with the polynucleotide of claim 24; and

(b) purifying said protein from the culture.

28. A protein produced according to the process of claim 27.

29. An isolated polynucleotide encoding the protein of claim 28.

30. The polynucleotide of claim 29, wherein the polynucleotide comprises the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818.

31. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) a fragment of the amino acid sequence of SEQ ID NO:6, the fragment comprising eight contiguous amino acids of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone dw665_4 deposited under accession number ATCC 98818;

the protein being substantially free from other mammalian proteins.

32. The protein of claim 31, wherein said protein comprises the amino acid sequence of SEQ ID NO:6.

33. A composition comprising the protein of claim 31 and a pharmaceutically acceptable carrier.

34. A method for preventing, treating, or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier; wherein the medical condition is selected from the group consisting of defects in cartilage, bone, or connective tissue formation, and damage to cartilage, bone, or connective tissue; and wherein the protein of the present invention comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:4 from amino acid 35 to amino acid 1036, SEQ ID NO:4 from amino acid 637 to amino acid 1036, SEQ ID NO:6, SEQ ID NO:6 from amino acid 28 to amino acid 457, and SEQ ID NO:6 from amino acid 29 to amino acid 457.

35. The method of claim 34 wherein the medical condition is selected from the group consisting of broken bones; congenital, trauma-induced, or oncologic-resection-induced craniofacial defects; periodontal disease; defects in the periodontal ligament or attachment apparatus; damage to the periodontal ligament or attachment apparatus; osteoporosis; burns; incisions; and ulcers.

Fig. 1A

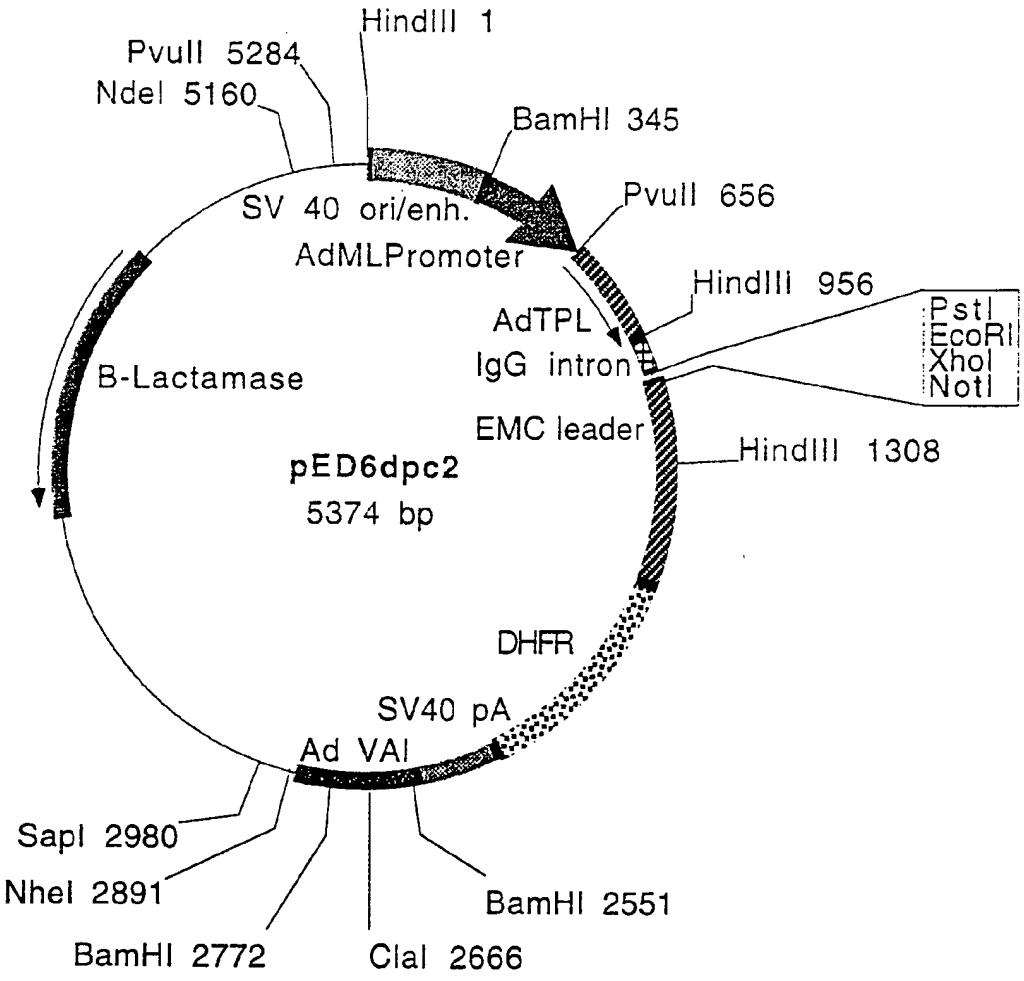
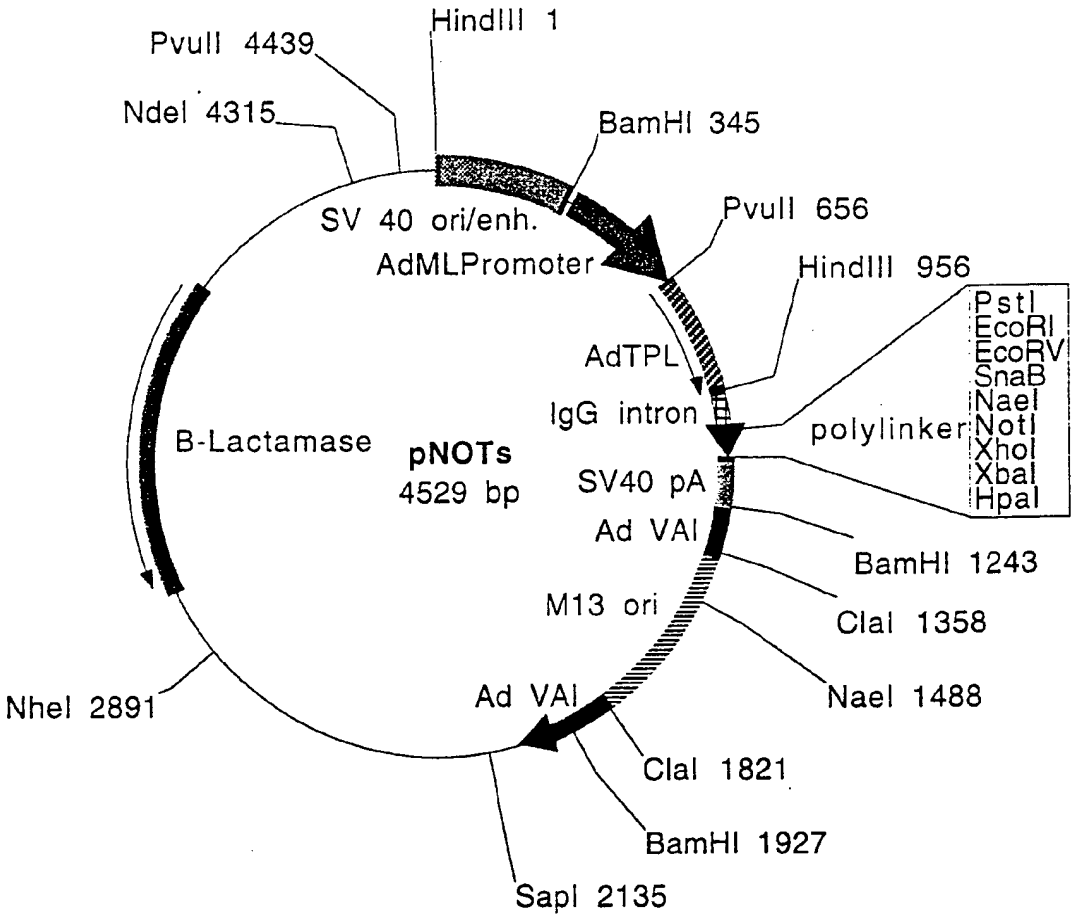

HindIII 1
PvuII 5284
NdeI 5160
BamHI 345
PvuII 656
SV 40 ori/enh.
AdMLPromoter
HindIII 956
AdTPL
IgG intron
PstI
EcoRI
XhoI
NotI
B-Lactamase
EMC leader
HindIII 1308
pED6dpc2
5374 bp
DHFR
SV40 pA
Ad VAI
SapI 2980
NheI 2891
BamHI 2772
ClaI 2666
BamHI 2551

Fig. 1B


HindIII 1
PvuII 4439
NdeI 4315
BamHI 345
SV 40 ori/enh.
AdMLPromoter
PvuII 656
HindIII 956
AdTPL
IgG intron
PstI
EcoRI
EcoRV
SnaB
NaeI
NotI
XhoI
XbaI
HpaI
polylinker
B-Lactamase
pNOTs
4529 bp
SV40 pA
Ad VAI
BamHI 1243
ClaI 1358
M13 ori
NaeI 1488
Ad VAI
NheI 2891
ClaI 1821
BamHI 1927
SapI 2135

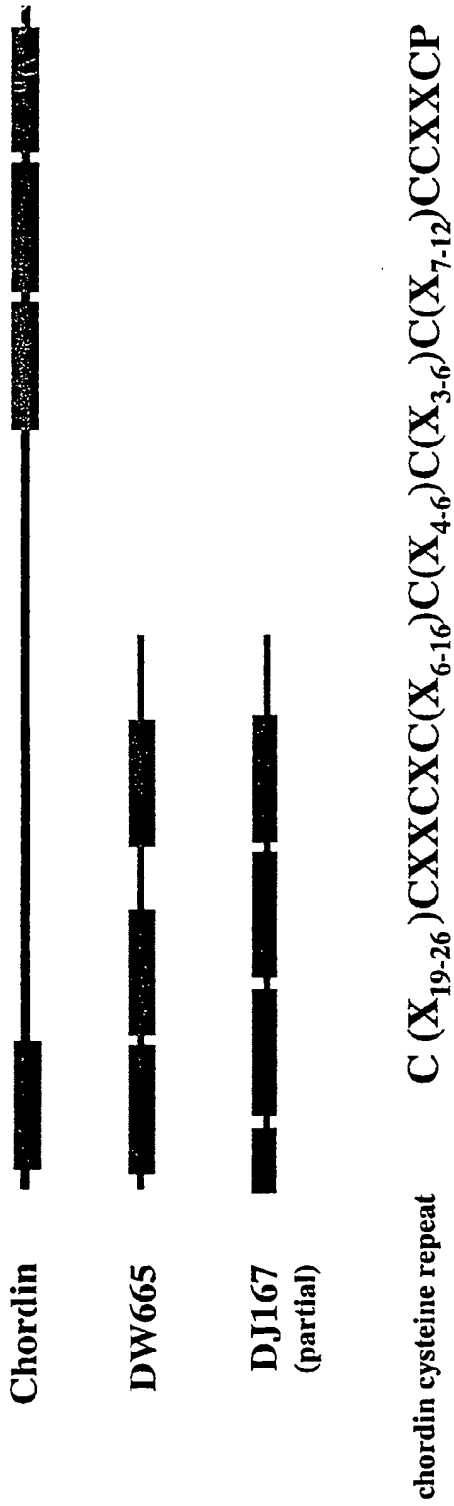

Chordin
DW665
DJ167
(partial)
chordin cysteine repeat
C (X19-26)CXXCXC(X6-16)C(X4-6)C(X3-6)C(X7-12)CCXXCP

Fig. 2

DJ167 tissue expression
Brain
Heart
Kidney
Spleen
Liver
Colon
Lung
Small Intestine
Muscle
Stomach
Testis
Placenta
Salivary Gland
Thyroid Gland
Adrenal Gland
Pancreas
Ovary
Uterus
Prostate
Skin
PBL
Bone Marrow
Fetal Brain
Fetal Liver
1000x
100x
10x
1x

Fig. 3

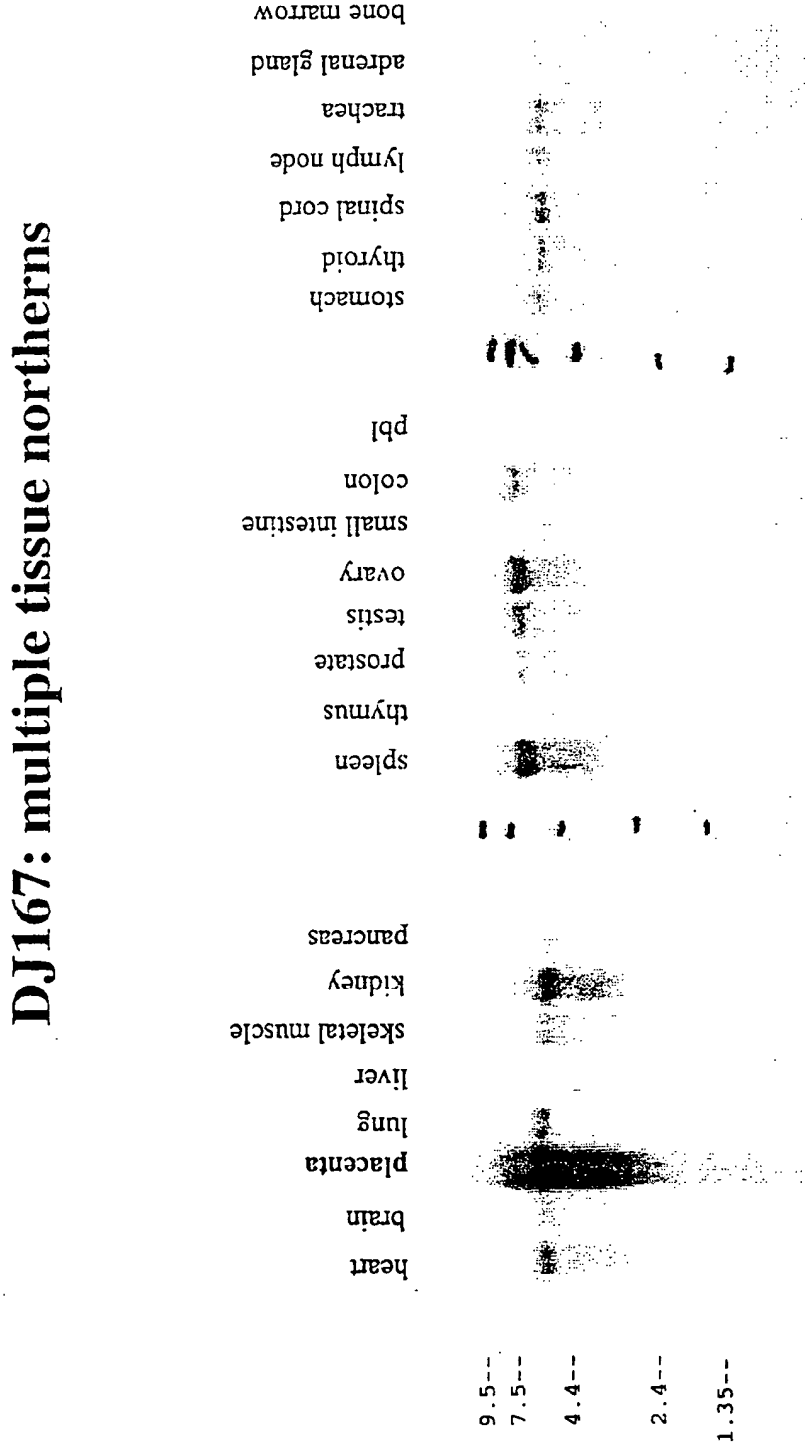

DJ167: multiple tissue northerns
9.5--
7.5--
4.4--
2.4--
1.35--
heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
spleen
thymus
prostate
testis
ovary
small intestine
colon
pbl
stomach
thyroid
spinal cord
lymph node
trachea
adrenal gland
bone marrow

Fig. 4

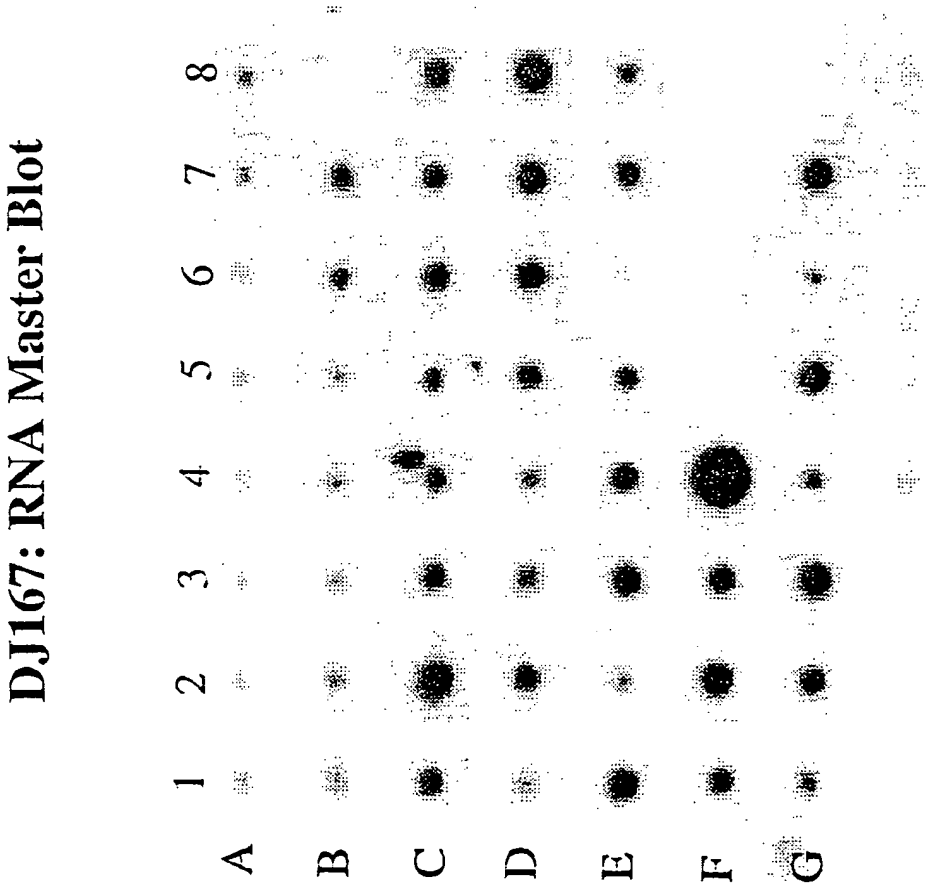


Fig. 5

DW665 tissue expression
Brain
Heart
Kidney
Spleen
Liver
Colon
Lung
Small Intestine
Muscle
Stomach
Testis
Placenta
Salivary Gland
Thyroid Gland
Adrenal Gland
Pancreas
Ovary
Uterus
Prostate
Skin
PBL
Bone Marrow
Fetal Brain
Fetal Liver
1000x
100x
10x
1x

Fig. 6

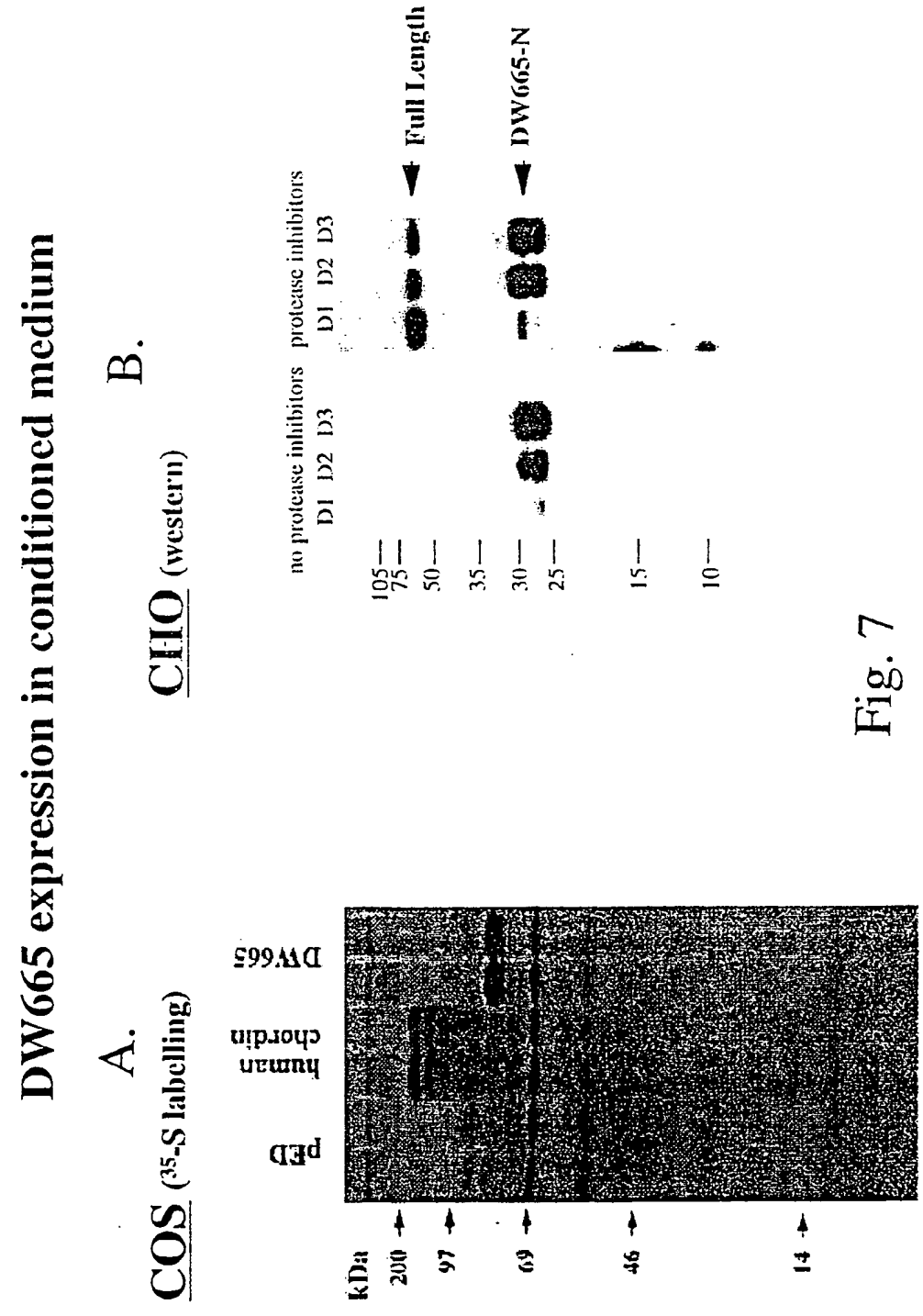

DW665 expression in conditioned medium
A.
COS (35-S labelling)
pED
human chordin
DW665
kDa
200
97
69
46
14
B.
CHO (western)
no protease inhibitors
D1 D2 D3
protease inhibitors
D1 D2 D3
105
75
50
35
30
25
15
10
Full Length
DW665-N

Fig. 7

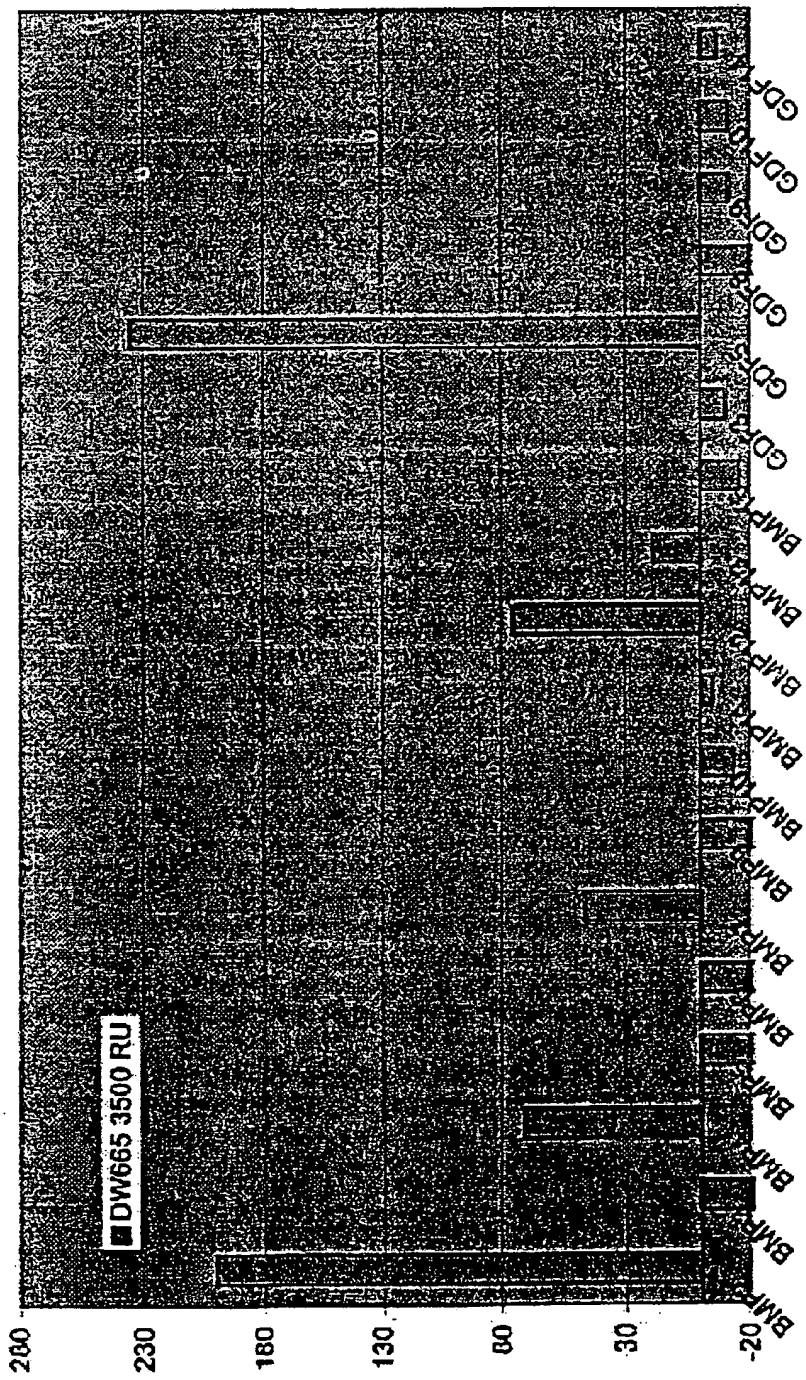

DW665 : spectrum of BMP binding
DW665 3500 RU
280
230
180
130
80
30
-20
GDF5
GDF6
GDF7
GDF8
GDF9
GDF10
GDF11
BMP2
BMP3
BMP4
BMP5
BMP6
BMP7
BMP8
BMP9
BMP10
BMP15

Fig. 8

DW665 induces a secondary axis in *Xenopus* embryos

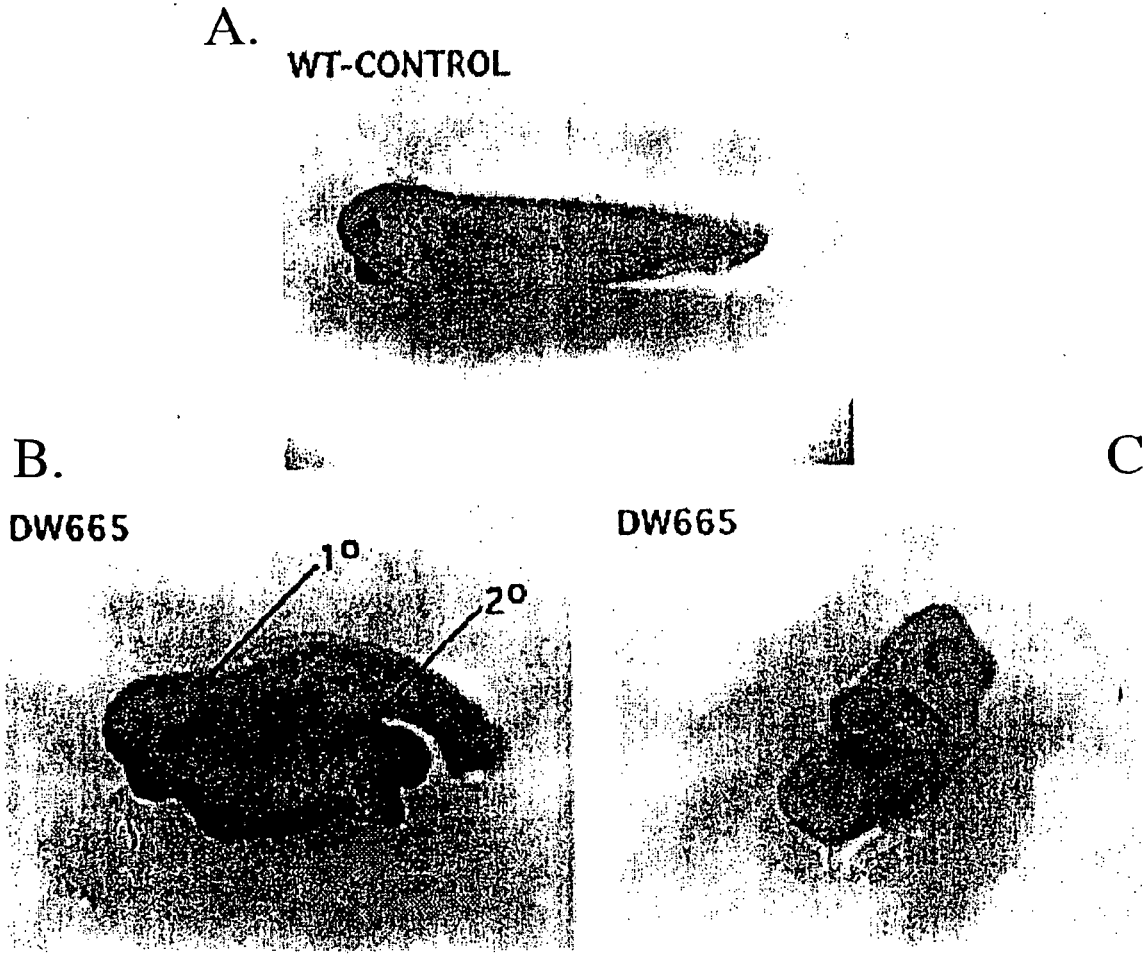


Fig. 9

10/10

SEQUENCE LISTING

<110> Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Collins-Racie, Lisa A.
Merberg, David
Treacy, Maurice
DiBlasio-Smith, Elizabeth
Widom, Angela
Genetics Institute, Inc.

<120> DJ167 AND DW665: HUMAN CHORDIN-RELATED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

<130> GI 6073A.DJ167DW665

<140>
<141>

<160> 9

<170> PatentIn Ver. 2.0

<210> 1
<211> 1456
<212> DNA
<213> Homo sapiens

<400> 1
cacggctgtc ttatctgcaa gtgcagagag gcctctgctt cagctgggcc acccatcctg 60
tcgggcactt gtctcaccgt ggatggtcat catcataaaa atgaggagag ctggcacgat 120
gggtgccggg aatgctactg tctcaatgga cgggaaatgt gtgccctgat cacctgcccg 180
gtgcctgcct gtggcaaccc caccattcac cctggacagt gctgcccatc atgtgcagat 240
gactttgtgg tgcagaagcc agagctcagt actccctcca tttgccacgc ccctggagga 300
gaatactttg tggaaggaga aacgtggaac attgactcct gtactcagtg cacctgccac 360
agcggacggg tgctgtgtga gacagaggtg tgcccaccgc tgctctgcca gaacccctca 420
cgcacccagg attcctgctg cccacagtgt acagatcaac cttttcggcc ttccttgtcc 480
cgcaataaca gcgtacctaa ttattgcaaa aatgatgaag gggatatatt cctggcagct 540
gagtcctgga agcctgacgt ttgtaccagc tgcatctgca ttgatagcgt aattagctgt 600
ttctctgagt cctgcccttc tgtatcctgt gaaagacctg tcttgagaaa aggccagtgt 660
tgtccctact gcatagaaga cacaattcca aagaaggtgg tgtgccactt cagtgggaag 720
gcctatgccg acgaggagcg gtgggacctt gacagctgca cccactgcta ctgcctgcag 780
ggccagaccc tctgctcgac cgtcagctgc ccccctctgc cctgtgttga gcccatcaac 840
gtggaaggaa gttgctgccc aatgtgtcca gaaatgtatg tcccagaacc aaccaatata 900
cccattgaga agacaaacca tcgaggagag gttgacctgg aggttccct gtggcccacg 960
cctagtgaaa atgatatcgt ccatctccct agagatatgg gtcacctcca ggtagattac 1020
agagataaca ggctgcaccc aagtgaagat tcttcactgg actccattgc ctcagttgtg 1080
gttcccataa ttatatgcct ctctattata atagcattcc tattcatcaa tcagaagaaa 1140
cagtggatac cactgctttg ctggtatcga acaccaacta agccttcttc cttaaataat 1200
cagctagtat ctgtggactg caagaaagga accagagtcc aggtggacag ttcccagaga 1260
atgctaagaa ttgcagaacc agatgcaaga ttcagtggct tctacagcat gcaaaaacag 1320
aaccatctac aggcagacaa tttctaccaa acagtgtgaa gaaaggcaac taggatgagg 1380
tttcaaaaga cggaagacga ctaaatctgc tctaaaaagt aaactagaat ttgtgcactt 1440
aaaaaaaaaa aaaaaa 1456

<210> 2
<211> 400
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Cys Ala Leu Ile Thr Cys Pro Val Pro Ala Cys Gly Asn Pro Thr
 1               5                  10                  15

Ile His Pro Gly Gln Cys Cys Pro Ser Cys Ala Asp Asp Phe Val Val
            20                  25                  30

Gln Lys Pro Glu Leu Ser Thr Pro Ser Ile Cys His Ala Pro Gly Gly
        35                  40                  45

Glu Tyr Phe Val Glu Gly Glu Thr Trp Asn Ile Asp Ser Cys Thr Gln
    50                  55                  60

Cys Thr Cys His Ser Gly Arg Val Leu Cys Glu Thr Glu Val Cys Pro
65                  70                  75                  80

Pro Leu Leu Cys Gln Asn Pro Ser Arg Thr Gln Asp Ser Cys Cys Pro
                85                  90                  95

Gln Cys Thr Asp Gln Pro Phe Arg Pro Ser Leu Ser Arg Asn Asn Ser
            100                 105                 110

Val Pro Asn Tyr Cys Lys Asn Asp Glu Gly Asp Ile Phe Leu Ala Ala
        115                 120                 125

Glu Ser Trp Lys Pro Asp Val Cys Thr Ser Cys Ile Cys Ile Asp Ser
    130                 135                 140

Val Ile Ser Cys Phe Ser Glu Ser Cys Pro Ser Val Ser Cys Glu Arg
145                 150                 155                 160

Pro Val Leu Arg Lys Gly Gln Cys Cys Pro Tyr Cys Ile Glu Asp Thr
                165                 170                 175

Ile Pro Lys Lys Val Val Cys His Phe Ser Gly Lys Ala Tyr Ala Asp
            180                 185                 190

Glu Glu Arg Trp Asp Leu Asp Ser Cys Thr His Cys Tyr Cys Leu Gln
        195                 200                 205

Gly Gln Thr Leu Cys Ser Thr Val Ser Cys Pro Pro Leu Pro Cys Val
    210                 215                 220

Glu Pro Ile Asn Val Glu Gly Ser Cys Cys Pro Met Cys Pro Glu Met
225                 230                 235                 240

Tyr Val Pro Glu Pro Thr Asn Ile Pro Ile Glu Lys Thr Asn His Arg
                245                 250                 255

Gly Glu Val Asp Leu Glu Val Pro Leu Trp Pro Thr Pro Ser Glu Asn
            260                 265                 270

Asp Ile Val His Leu Pro Arg Asp Met Gly His Leu Gln Val Asp Tyr
        275                 280                 285

Arg Asp Asn Arg Leu His Pro Ser Glu Asp Ser Ser Leu Asp Ser Ile
    290                 295                 300

Ala Ser Val Val Val Pro Ile Ile Ile Cys Leu Ser Ile Ile Ile Ala
305                 310                 315                 320
```

Phe Leu Phe Ile Asn Gln Lys Lys Gln Trp Ile Pro Leu Leu Cys Trp
325 330 335

Tyr Arg Thr Pro Thr Lys Pro Ser Ser Leu Asn Asn Gln Leu Val Ser
340 345 350

Val Asp Cys Lys Lys Gly Thr Arg Val Gln Val Asp Ser Ser Gln Arg
355 360 365

Met Leu Arg Ile Ala Glu Pro Asp Ala Arg Phe Ser Gly Phe Tyr Ser
370 375 380

Met Gln Lys Gln Asn His Leu Gln Ala Asp Asn Phe Tyr Gln Thr Val
385 390 395 400

<210> 3
<211> 4723
<212> DNA
<213> Homo sapiens

<400> 3
ggccttcatg gcctattttt ttttttttt aaatgataca acttaatttt attaggacaa 60
ggctggtggg cactggagtg gcaccttcag ggccaggaga ggcactgggg aggggtcaca 120
ggatgctact cgggcaccta gaagccacag ctgccctcca cagagcggca ctgcaccatg 180
cgcaggaatg tctcgacctt gtccatgtcc ttcctgaagc agtagagcag cccgtagttc 240
ttgagcagtg cgtcatggtt gtgcgagttt gtgtcaaact tgctgtaggt ctgcttgagg 300
atctgcccag tccggcggct gccgtcttcc agcctcccca tcagcgtttg gatgccttcc 360
tctaggtcct ttaggaggtg atagtcatcg ttgtccgagg tgtcatacac caggttgttg 420
gcgaacatac tcctgaggaa ccgcacgggc tccagccacg actcgatgag cagcagggag 480
atgcggagca gctctagatt ggatttctgt tgcgtttcct ccatgttgga gggtgtcgga 540
atagagtctg agaagcagaa ggaggtctgg gagtcatgca ggaatgaata cttctggtcc 600
tttgggatat aggtttcttc aaactcctgg taggtgtcaa tggccagctg gtgcgcgcga 660
tgggcttgga gcatagcgtg gtcaaaaagc ctggataacg gaacggtttg gacggcacca 720
gcctcttgaa gccagggcag gcagagcagg gcaaaagcca ggagcaggga cgtccgggag 780
cctggagcca ttgccactag gtgagctgtc cacaggaccc tgagtggttc ggggagttcg 840
gccttcatgg cctaggagcg gcgcaggagt gaggcgagcg gggcgcgcgg agcggacgcc 900
gcggatcttg tgctgcgcca ccgcgcccac tcggcagctc gggaggcggg gaccggcccg 960
gaggctgcgc cgctgcgggg ccggccgact cggaggagga gagggaggag gcgccgccgg 1020
cccgggctgg agccgagcgc agcagccacc gccgccgccg cgccagaagt ttgggttgaa 1080
ccggagctgc cgggaggaaa ctttttttctt ttttccccct ccctcccggg aggaggagga 1140
ggaggaggag gggaagctgc cgccggcgcc aaggctcgtg ggctcggggt cggcgcggcc 1200
cgcagaaggg gcgggggcct cgccccgcga ggggaggcgc gccccggggg ccccgagagg 1260
ggcggtgagg accgcgggct gctggtgcgg cggcggcggc ggcgcgtgtg ccccgcgcag 1320
gggagggcgc ccgccccgct cccggcccgg ctgcgaggag gaggcggcgg cggcgcagga 1380
ggatgtactt ggtggcgggg gacaggggat tggccggctg cgggcacctc ctggtctcgc 1440
tgctggggct gctgctgctg ctggcgcgct ccggcacccg ggcgctggtc tgcctgccct 1500
gtgacgagtc caagtgcgag gagcccagga actgcccggg gagcatcgtg cagggcgtct 1560
gcggctgctg ctacacgtgc gccagccaga ggaacgagag ctgcggcggc accttcggga 1620
tttacggaac ctgcgaccgg gggctgcgtt gtgtcatccg ccccccgctc aatggcgact 1680
ccctcaccga gtacgaagcg ggcgtttgcg aagatgagaa ctggactgat gaccaactgc 1740
ttggttttaa accatgcaat gaaaacctta ttgctggctg caatataatc aatgggaaat 1800
gtgaatgtaa caccattcga acctgcagca atccctttga gtttccaagt caggatatgt 1860
gcctttcagc tttaaagaga attgaagaag agaagccaga ttgctccaag gcccgctgtg 1920
aagtccagtt ctctccacgt tgtcctgaag attctgttct gatcgagggt tatgctcctc 1980
ctggggagtg ctgtccctta cccagccgct gcgtgtgcaa ccccgcaggc tgtctgcgca 2040
aagtctgcca gccgggaaac ctgaacatac tagtgtcaaa agcctcaggg aagccgggag 2100
agtgctgtga cctctatgag tgcaaaccag ttttcggcgt ggactgcagg actgtggaat 2160
gccctcctgt tcagcagacc gcgtgtcccc cggacagcta tgaaactcaa gtcagactaa 2220

```
ctgcagatgg ttgctgtact ttgccaacaa gatgcgagtg tctctctggc ttatgtggtt 2280
tccccgtgtg tgaggtggga tccactcccc gcatagtctc tcgtggcgat gggacacctg 2340
gaaagtgctg tgatgtcttt gaatgtgtta atgatacaaa gccagcctgc gtatttaaca 2400
atgtggaata ttatgatgga gacatgtttc gaatggacaa ctgtcggttc tgtcgatgcc 2460
aaggggggcgt tgccatctgc ttcactgccc agtgtggtga gataaactgc gagaggtact 2520
acgtgcccga aggagagtgc tgcccagtgt gtgaagatcc agtgtatcct tttaataatc 2580
ccgctggctg ctatgccaat ggcctgatcc ttgcccacgg agaccggtgg cgggaagacg 2640
actgcacatt ctgccagtgc gtcaacggtg aacgccactg cgttgcgacc gtctgcggac 2700
agacctgcac aaaccctgtg aaagtgcctg gggagtgttg ccctgtgtgc gaagaaccaa 2760
ccatcatcac agttgatcca cctgcatgtg gggagttatc aaactgcact ctgacaggga 2820
aggactgcat taatggtttc aaacgcgatc acaatggttg tcggaccctgt cagtgcataa 2880
acaccgagga actatgttca gaacgtaaac aaggctgcac cttgaactgt cccttcggtt 2940
tccttactga tgcccaaaac tgtgagatct gtgagtgccg cccaaggccc aagaagtgca 3000
gacccataat ctgtgacaag tattgtccac ttggattgct gaagaataag cacggctgtg 3060
acatctgtcg ctgtaagaaa tgtccagagc tctcatgcag taagatctgc cccttgggtt 3120
tccagcagga cagtcacggc tgtcttatct gcaagtgcag agaggcctct gcttcagctg 3180
ggccacccat cctgtcgggc acttgtctca ccgtggatgg tcatcatcat aaaaatgagg 3240
agagctggca cgatgggtgc cgggaatgct actgtctcaa tggacgggaa atgtgtgccc 3300
tgatcacctg cccggtgcct gcctgtggca accccaccat tcaccctgga cagtgctgcc 3360
catcatgtgc agatgacttt gtggtgcaga agccagagct cagtactccc tccatttgcc 3420
acgcccctgg aggagaatac tttgtggaag gagaaacgtg gaacattgac tcctgtactc 3480
agtgcacctg ccacagcgga cgggtgctgt gtgagacaga ggtgtgccca ccgctgctct 3540
gccagaaccc ctcacgcacc caggattcct gctgcccaca gtgtacagat caacctttc 3600
ggccttcctt gtcccgcaat aacagcgtac ctaattactg caaaaatgat gaaggggata 3660
tattcctggc agctgagtcc tggaagcctg acgtttgtac cagctgcatc tgcattgata 3720
gcgtaattag ctgtttctct gagtcctgcc cttctgtatc ctgtgaaaga cctgtcttga 3780
gaaaaggcca gtgttgtccc tactgcatag aagacacaat tccaaagaag gtggtgtgcc 3840
acttcagtgg gaaggcctat gccgacgagg agcggtggga ccttgacagc tgcacccact 3900
gctactgcct gcagggccag accctctgct cgaccgtcag ctgccccct ctgccctgtg 3960
ttgagcccat caacgtggaa ggaagttgct gcccaatgtg tccagaaatg tatgtcccag 4020
aaccaaccaa tatacccatt gagaagacaa accatcgagg agaggttgac ctggaggttc 4080
ccctgtggcc cacgcctagt gaaaatgata tcgtccatct ccctagagat atgggtcacc 4140
tccaggtaga ttacagagat aacaggctgc acccaagtga agattcttca ctggactcca 4200
ttgcctcagt tgtggttccc ataattatat gcctctctat tataatagca ttcctattca 4260
tcaatcagaa gaaacagtgg ataccactgc tttgctggta tcgaacacca actaagcctt 4320
cttccttaaa taatcagtta gtatctgtgg actgcaagaa aggaaccaga gtccaggtgg 4380
acagttccca gagaatgcta agaattgcag aaccagatgc aagattcagt ggcttctaca 4440
gcatgcaaaa acagaaccat ctacaggcag acaatttcta ccaaacagtg tgaagaaagg 4500
caactaggat gaggtttcaa aagacggaag acgactaaat ctgctctaaa aagtaaacta 4560
gaatttgtgc acttgcttag tggattgtat tggattgtga cttgatgtac agcgctaaga 4620
ccttactggg atgggctctg tctacagcaa tgtgcagaac aagcattccc cctcaaacct 4680
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa                    4723
```

<210> 4
<211> 1036
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Tyr Leu Val Ala Gly Asp Arg Gly Leu Ala Gly Cys Gly His Leu
  1               5                  10                  15

Leu Val Ser Leu Leu Gly Leu Leu Leu Leu Leu Ala Arg Ser Gly Thr
             20                  25                  30

Arg Ala Leu Val Cys Leu Pro Cys Asp Glu Ser Lys Cys Glu Glu Pro
         35                  40                  45

Arg Asn Cys Pro Gly Ser Ile Val Gln Gly Val Cys Gly Cys Cys Tyr
     50                  55                  60
```

```
Thr Cys Ala Ser Gln Arg Asn Glu Ser Cys Gly Gly Thr Phe Gly Ile
 65                  70                  75                  80

Tyr Gly Thr Cys Asp Arg Gly Leu Arg Cys Val Ile Arg Pro Pro Leu
                85                  90                  95

Asn Gly Asp Ser Leu Thr Glu Tyr Glu Ala Gly Val Cys Glu Asp Glu
            100                 105                 110

Asn Trp Thr Asp Asp Gln Leu Leu Gly Phe Lys Pro Cys Asn Glu Asn
        115                 120                 125

Leu Ile Ala Gly Cys Asn Ile Ile Asn Gly Lys Cys Glu Cys Asn Thr
    130                 135                 140

Ile Arg Thr Cys Ser Asn Pro Phe Glu Phe Pro Ser Gln Asp Met Cys
145                 150                 155                 160

Leu Ser Ala Leu Lys Arg Ile Glu Glu Glu Lys Pro Asp Cys Ser Lys
                165                 170                 175

Ala Arg Cys Glu Val Gln Phe Ser Pro Arg Cys Pro Glu Asp Ser Val
            180                 185                 190

Leu Ile Glu Gly Tyr Ala Pro Pro Gly Glu Cys Cys Pro Leu Pro Ser
        195                 200                 205

Arg Cys Val Cys Asn Pro Ala Gly Cys Leu Arg Lys Val Cys Gln Pro
    210                 215                 220

Gly Asn Leu Asn Ile Leu Val Ser Lys Ala Ser Gly Lys Pro Gly Glu
225                 230                 235                 240

Cys Cys Asp Leu Tyr Glu Cys Lys Pro Val Phe Gly Val Asp Cys Arg
                245                 250                 255

Thr Val Glu Cys Pro Pro Val Gln Gln Thr Ala Cys Pro Pro Asp Ser
            260                 265                 270

Tyr Glu Thr Gln Val Arg Leu Thr Ala Asp Gly Cys Cys Thr Leu Pro
        275                 280                 285

Thr Arg Cys Glu Cys Leu Ser Gly Leu Cys Gly Phe Pro Val Cys Glu
    290                 295                 300

Val Gly Ser Thr Pro Arg Ile Val Ser Arg Gly Asp Gly Thr Pro Gly
305                 310                 315                 320

Lys Cys Cys Asp Val Phe Glu Cys Val Asn Asp Thr Lys Pro Ala Cys
                325                 330                 335

Val Phe Asn Asn Val Glu Tyr Tyr Asp Gly Asp Met Phe Arg Met Asp
            340                 345                 350

Asn Cys Arg Phe Cys Arg Cys Gln Gly Gly Val Ala Ile Cys Phe Thr
        355                 360                 365

Ala Gln Cys Gly Glu Ile Asn Cys Glu Arg Tyr Tyr Val Pro Glu Gly
    370                 375                 380
```

```
Glu Cys Cys Pro Val Cys Glu Asp Pro Val Tyr Pro Phe Asn Asn Pro
385                 390                 395                 400

Ala Gly Cys Tyr Ala Asn Gly Leu Ile Leu Ala His Gly Asp Arg Trp
                405                 410                 415

Arg Glu Asp Asp Cys Thr Phe Cys Gln Cys Val Asn Gly Glu Arg His
            420                 425                 430

Cys Val Ala Thr Val Cys Gly Gln Thr Cys Thr Asn Pro Val Lys Val
        435                 440                 445

Pro Gly Glu Cys Cys Pro Val Cys Glu Glu Pro Thr Ile Ile Thr Val
    450                 455                 460

Asp Pro Pro Ala Cys Gly Glu Leu Ser Asn Cys Thr Leu Thr Gly Lys
465                 470                 475                 480

Asp Cys Ile Asn Gly Phe Lys Arg Asp His Asn Gly Cys Arg Thr Cys
                485                 490                 495

Gln Cys Ile Asn Thr Glu Glu Leu Cys Ser Glu Arg Lys Gln Gly Cys
            500                 505                 510

Thr Leu Asn Cys Pro Phe Gly Phe Leu Thr Asp Ala Gln Asn Cys Glu
        515                 520                 525

Ile Cys Glu Cys Arg Pro Arg Pro Lys Lys Cys Arg Pro Ile Ile Cys
    530                 535                 540

Asp Lys Tyr Cys Pro Leu Gly Leu Leu Lys Asn Lys His Gly Cys Asp
545                 550                 555                 560

Ile Cys Arg Cys Lys Lys Cys Pro Glu Leu Ser Cys Ser Lys Ile Cys
                565                 570                 575

Pro Leu Gly Phe Gln Gln Asp Ser His Gly Cys Leu Ile Cys Lys Cys
            580                 585                 590

Arg Glu Ala Ser Ala Ser Ala Gly Pro Pro Ile Leu Ser Gly Thr Cys
        595                 600                 605

Leu Thr Val Asp Gly His His His Lys Asn Glu Glu Ser Trp His Asp
    610                 615                 620

Gly Cys Arg Glu Cys Tyr Cys Leu Asn Gly Arg Glu Met Cys Ala Leu
625                 630                 635                 640

Ile Thr Cys Pro Val Pro Ala Cys Gly Asn Pro Thr Ile His Pro Gly
                645                 650                 655

Gln Cys Cys Pro Ser Cys Ala Asp Asp Phe Val Val Gln Lys Pro Glu
            660                 665                 670

Leu Ser Thr Pro Ser Ile Cys His Ala Pro Gly Gly Glu Tyr Phe Val
        675                 680                 685

Glu Gly Glu Thr Trp Asn Ile Asp Ser Cys Thr Gln Cys Thr Cys His
    690                 695                 700
```

```
Ser Gly Arg Val Leu Cys Glu Thr Glu Val Cys Pro Pro Leu Leu Cys
705                 710                 715                 720

Gln Asn Pro Ser Arg Thr Gln Asp Ser Cys Cys Pro Gln Cys Thr Asp
                725                 730                 735

Gln Pro Phe Arg Pro Ser Leu Ser Arg Asn Asn Ser Val Pro Asn Tyr
            740                 745                 750

Cys Lys Asn Asp Glu Gly Asp Ile Phe Leu Ala Ala Glu Ser Trp Lys
        755                 760                 765

Pro Asp Val Cys Thr Ser Cys Ile Cys Ile Asp Ser Val Ile Ser Cys
    770                 775                 780

Phe Ser Glu Ser Cys Pro Ser Val Ser Cys Glu Arg Pro Val Leu Arg
785                 790                 795                 800

Lys Gly Gln Cys Cys Pro Tyr Cys Ile Glu Asp Thr Ile Pro Lys Lys
                805                 810                 815

Val Val Cys His Phe Ser Gly Lys Ala Tyr Ala Asp Glu Glu Arg Trp
            820                 825                 830

Asp Leu Asp Ser Cys Thr His Cys Tyr Cys Leu Gln Gly Gln Thr Leu
        835                 840                 845

Cys Ser Thr Val Ser Cys Pro Pro Leu Pro Cys Val Glu Pro Ile Asn
    850                 855                 860

Val Glu Gly Ser Cys Cys Pro Met Cys Pro Glu Met Tyr Val Pro Glu
865                 870                 875                 880

Pro Thr Asn Ile Pro Ile Glu Lys Thr Asn His Arg Gly Glu Val Asp
                885                 890                 895

Leu Glu Val Pro Leu Trp Pro Thr Pro Ser Glu Asn Asp Ile Val His
            900                 905                 910

Leu Pro Arg Asp Met Gly His Leu Gln Val Asp Tyr Arg Asp Asn Arg
        915                 920                 925

Leu His Pro Ser Glu Asp Ser Ser Leu Asp Ser Ile Ala Ser Val Val
    930                 935                 940

Val Pro Ile Ile Ile Cys Leu Ser Ile Ile Ile Ala Phe Leu Phe Ile
945                 950                 955                 960

Asn Gln Lys Lys Gln Trp Ile Pro Leu Leu Cys Trp Tyr Arg Thr Pro
                965                 970                 975

Thr Lys Pro Ser Ser Leu Asn Asn Gln Leu Val Ser Val Asp Cys Lys
            980                 985                 990

Lys Gly Thr Arg Val Gln Val Asp Ser Ser Gln Arg Met Leu Arg Ile
        995                1000                1005

Ala Glu Pro Asp Ala Arg Phe Ser Gly Phe Tyr Ser Met Gln Lys Gln
   1010                1015                1020
```

Asn His Leu Gln Ala Asp Asn Phe Tyr Gln Thr Val
1025 1030 1035

<210> 5
<211> 3861
<212> DNA
<213> Homo sapiens

<400> 5
gtgcacgcgt ggcagacgga gaaggccagt gcccagcttg aaggttctgt cacctttgc 60
agtggtccaa atgagaaaaa agtggaaaat gggaggcatg aaatacatct tttcgttgtt 120
gttctttctt ttgctagaag gaggcaaaac agagcaagta aaacattcag agacatattg 180
catgtttcaa gacaagaagt acagagtggg tgagagatgg catccttacc tggaaccttа 240
tgggttggtt tactgcgtga actgcatctg ctcagagaat gggaatgtgc tttgcagccg 300
agtcagatgt ccaaatgttc attgcctttc tcctgtgcat attcctcatc tgtgctgccc 360
tcgctgccca gactccttac ccccagtgaa caataaggtg accagcaagt cttgcgagta 420
caatgggaca acttaccaac atggagagct gttcgtagct gaagggctct ttcagaatcg 480
gcaacccaat caatgcaccc agtgcagctg ttcggaggga aacgtgtatt gtggtctcaa 540
gacttgcccc aaattaacct gtgccttccc agtctctgtt ccagattcct gctgccgggt 600
atgcagagga gatggagaac tgtcatggga acattctgat ggtgatatct tccggcaacc 660
tgccaacaga gaagcaagac attcttacca ccgctctcac tatgatcctc caccaagccg 720
acaggctgga ggtctgtccc gctttcctgg ggccagaagt caccggggag ctcttatgga 780
ttcccagcaa gcatcaggaa ccattgtgca aattgtcatc aataacaaac acaagcatgg 840
acaagtgtgt gtttccaatg gaaagaccta ttctcatggc gagtcctggc acccaaacct 900
ccgggcattt ggcattgtgg agtgtgtgct atgtacttgt aatgtcacca agcaagagtg 960
taagaaaatc cactgcccca atcgataccc ctgcaagtat cctcaaaaaa tagacggaaa 1020
gtgctgcaag gtgtgtccag gtaaaaaagc aaaagaagaa cttccaggcc aaagctttga 1080
caataaaggc tacttctgcg gggaagaaac gatgcctgtg tatgagtctg tattcatgga 1140
ggatggggag acaaccagaa aaatagcact ggagactgag agaccacctc aggtagaggt 1200
ccacgtttgg actattcgaa agggcattct ccagcacttc catattgaga agatctccaa 1260
gaggatgttt gaggagcttc ctcacttcaa gctggtgacc agaacaaccc tgagccagtg 1320
gaagatcttc accgaaggag aagctcagat cagccagatg tgttcaagtc gtgtatgcag 1380
aacagagctt gaagatttag tcaaggtttt gtacctggag agatctgaaa agggccactg 1440
ttaggcaaga cagacagtat tggatagggt aaagcaagaa aactcaagct gcagctggac 1500
tgcaggctta ttttgcttaa gtcaacagtg ccctaaaact ccaaactcaa atgcagtcaa 1560
ttattcacgc catgcacagc ataatttgct cctttgtgtg gagtggtgtg tcagcccttg 1620
aacatctcct ccaaagagac tagaagagtc ttaaattata tgtgggagga ggagggatag 1680
aacatcacaa cactgctcta gtttcttgga gaatcacatt tctttacagg ttaaagacaa 1740
acaagacccc agggttttta tctagaaagt tattcaagtg aaagaaagag aagggaattg 1800
cttagtagga gttctgcagt atagaacaat tacttgtatg aaattatacc tttgaatttt 1860
agaatgtcat gtgttctttt aaaaaaatta gctccccatc ctcctcctc actcctctcc 1920
tccctccttc tctctctctc tctctctccc tctctcacag acacacacac acacacacac 1980
acacacacgc acgtccacac tcacattaaa ctaaagcttt atttgaagca aagctagcca 2040
aaattctacg ttacttttcc cttgactgga tcccaagtag cttggaagtt tttgtgccca 2100
ggagagtaaa taactgtgaa caagaggctc tgcccttagg tctttgtggc tgtttaagtc 2160
accaacaata gagtcagggt aaagaataaa aacactttca tagcctcatt cattcactta 2220
gaagtggtaa taatttttcc ctaatgatac cacttttctt ttcccctgt acctatggga 2280
cttccagaaa gaagttaaat tgagtaaaat catcagaaac tgaatccatg taagaaaaaa 2340
taattgttga agaaagaagt tgatagaatt caaaaaggcc atctttttgc tttcacatca 2400
ataaaattta ccaagtaata gatcagtact cactaatatt tttgagacca tagttgtctg 2460
gtcagaaaaa ttatattaaa ttagtaaatt ctagaagctc tttaaaaggg aagttttcct 2520
tcttctccaa ttataggagt tgatttttac tttgcaaagt ggctcggtcc tcatgagcat 2580
ctgcatgttg actcttcagt taagaaaatt gttgttcatt tagggaggtg gatattctga 2640
tgaagatctt tatcctaaac cttcctacta tccttgtctt attcatcaag cagatatttt 2700
agtcaagaat tccagagaag gctgctccta aaatgtctac ttgcagccca ataccagagc 2760
ataaactatc catttctgggg tctggcttta gaaatcatct ttgtgggaag acctaattct 2820
tcacagcaag gatctcaggc atgccttcta gatttgttcc ctctgagggg caggaatgaa 2880
ctgtagaaat gttttaagga cccagaaacc ccatatgtct cattccatga ctataggtga 2940

gagaattctt tcctaagagg gtttgatacc aataggggaa aatgtaaaat gttcagtctt 3000
tatgacaacc tggcataaag gagtcaattc ttatgaaaga gacacaaggg ccttatggcc 3060
agggtttctt gggacaagac tctcaccagc acatcacaca cgttctcctt ggaagagaga 3120
agcagtacat cccggttgag aggtcacaaa gcattagtgt gtgtgtgtgt gtgtgtgtgt 3180
gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtatgt ggtaaagggg ggaaggtgtt 3240
atgcggctgc tccctccgtc ccagaggtgg cagtgattcc ataatgtgga gactagtaac 3300
tagatcctaa ggcaaagagg tgtttctcct tttggatgat tcatcccaaa gccttcccac 3360
ccaggtgttc tctgaaagct tagccttaag agaacacgca gagagtttcc ctagatatac 3420
tcctgcctcc aggtgctggg acacaccttt gcaaaatgct gtgggaagca ggagctgggg 3480
agctgtgtta agtcaaagta gaaaccctcc agtgtttggt gttgtgtaga gaataggaca 3540
tagggtaaag aggccaagct gcctgtagtt agtagagaag aatggatgtg gttcttcttg 3600
tgtatttatt tgtatcataa acacttggaa caacaaagac cataagcatc atttagcagt 3660
tgtagccatt ttctagttaa ctcatgtaaa caagtaagag taacataaca gtattaccct 3720
ttcactgttc tcacaggaca tgtacctaat tatggtactt atttatgtag tcactgtatt 3780
tctggatttt taaattaata aaaaagttaa ttttgaaaaa tcaaaaaaaa aaaaaaaaaa 3840
aaaaaaaaaa aaaaaaaaaa a 3861

<210> 6
<211> 457
<212> PRT
<213> Homo sapiens

<400> 6
Met Arg Lys Lys Trp Lys Met Gly Gly Met Lys Tyr Ile Phe Ser Leu
1 5 10 15

Leu Phe Phe Leu Leu Leu Glu Gly Gly Lys Thr Glu Gln Val Lys His
20 25 30

Ser Glu Thr Tyr Cys Met Phe Gln Asp Lys Lys Tyr Arg Val Gly Glu
35 40 45

Arg Trp His Pro Tyr Leu Glu Pro Tyr Gly Leu Val Tyr Cys Val Asn
50 55 60

Cys Ile Cys Ser Glu Asn Gly Asn Val Leu Cys Ser Arg Val Arg Cys
65 70 75 80

Pro Asn Val His Cys Leu Ser Pro Val His Ile Pro His Leu Cys Cys
85 90 95

Pro Arg Cys Pro Asp Ser Leu Pro Pro Val Asn Asn Lys Val Thr Ser
100 105 110

Lys Ser Cys Glu Tyr Asn Gly Thr Thr Tyr Gln His Gly Glu Leu Phe
115 120 125

Val Ala Glu Gly Leu Phe Gln Asn Arg Gln Pro Asn Gln Cys Thr Gln
130 135 140

Cys Ser Cys Ser Glu Gly Asn Val Tyr Cys Gly Leu Lys Thr Cys Pro
145 150 155 160

Lys Leu Thr Cys Ala Phe Pro Val Ser Val Pro Asp Ser Cys Cys Arg
165 170 175

Val Cys Arg Gly Asp Gly Glu Leu Ser Trp Glu His Ser Asp Gly Asp
180 185 190

Ile Phe Arg Gln Pro Ala Asn Arg Glu Ala Arg His Ser Tyr His Arg

```
       195                 200                 205

Ser His Tyr Asp Pro Pro Pro Ser Arg Gln Ala Gly Gly Leu Ser Arg
    210                 215                 220

Phe Pro Gly Ala Arg Ser His Arg Gly Ala Leu Met Asp Ser Gln Gln
225                 230                 235                 240

Ala Ser Gly Thr Ile Val Gln Ile Val Ile Asn Asn Lys His Lys His
                245                 250                 255

Gly Gln Val Cys Val Ser Asn Gly Lys Thr Tyr Ser His Gly Glu Ser
            260                 265                 270

Trp His Pro Asn Leu Arg Ala Phe Gly Ile Val Glu Cys Val Leu Cys
        275                 280                 285

Thr Cys Asn Val Thr Lys Gln Glu Cys Lys Lys Ile His Cys Pro Asn
    290                 295                 300

Arg Tyr Pro Cys Lys Tyr Pro Gln Lys Ile Asp Gly Lys Cys Cys Lys
305                 310                 315                 320

Val Cys Pro Gly Lys Lys Ala Lys Glu Glu Leu Pro Gly Gln Ser Phe
                325                 330                 335

Asp Asn Lys Gly Tyr Phe Cys Gly Glu Glu Thr Met Pro Val Tyr Glu
            340                 345                 350

Ser Val Phe Met Glu Asp Gly Glu Thr Thr Arg Lys Ile Ala Leu Glu
        355                 360                 365

Thr Glu Arg Pro Pro Gln Val Glu Val His Val Trp Thr Ile Arg Lys
    370                 375                 380

Gly Ile Leu Gln His Phe His Ile Glu Lys Ile Ser Lys Arg Met Phe
385                 390                 395                 400

Glu Glu Leu Pro His Phe Lys Leu Val Thr Arg Thr Thr Leu Ser Gln
                405                 410                 415

Trp Lys Ile Phe Thr Glu Gly Glu Ala Gln Ile Ser Gln Met Cys Ser
            420                 425                 430

Ser Arg Val Cys Arg Thr Glu Leu Glu Asp Leu Val Lys Val Leu Tyr
        435                 440                 445

Leu Glu Arg Ser Glu Lys Gly His Cys
    450                 455


<210> 7
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<220>
```

<221> misc_feature
<222> (2)
<223> biotinylated phosphoaramidite residue

<400> 7
angttgatgg gctcaacaca gggcagagg 29

<210> 8
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide

<220>
<221> misc_feature
<222> (2)
<223> biotinylated phosphoaramidite residue

<400> 8
anggatgcca tctctcaccc actctgtac 29

<210> 9
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<221> NON_CONS
<222> (1)..(2)
<223> These cysteine residues are separated by 19 to 26 intervening amino acid residues

<220>
<221> UNSURE
<222> (3)..(4)

<220>
<221> UNSURE
<222> (6)

<220>
<221> NON_CONS
<222> (7)..(8)
<223> These cysteine residues are separated by 6 to 16 intervening amino acid residues

<220>
<221> NON_CONS
<222> (8)..(9)
<223> These cysteine residues are separated by 4 to 6 intervening amino acid residues

<220>
<221> NON_CONS
<222> (9)..(10)
<223> These cysteine residues are separated by 3 to 6 intervening amino acid residues

<220>
<221> NON_CONS
<222> (10)..(11)
<223> These cysteine residues are separated by 7 to 12
intervening amino acid residues

<220>
<221> UNSURE
<222> (13)..(14)

<220>
<223> consensus chordin cysteine repeat

<400> 9
Cys Cys Xaa Xaa Cys Xaa Cys Cys Cys Cys Cys Cys Xaa Xaa Cys Pro
1 5 10 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/18117

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1, 23.5; 530/300, 350; 435/69.1, 320.1, 325, 252.3, 254.11; 514/2, 12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, MEDLINE
search terms: dj167?, kenneth jacobs

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	WO 98/45436 A2 (GENETICS INSTITUTE, INC.) 15 October 1998, especially SEQ ID NO:1236 on page 488 and in claim 1.	1
X, P	GenBank Database, Accession Number AA964264, BONALDO et al. 'Normalization and subtraction: two approaches to facilitate gene discovery', abstract, 04 July 1999, Genome Res. 6(9), 791-806, see entire reference.	1-3
X	GenBank Database, Accession Number T98917, HILLIER et al. 'The WashU-Merck EST Project', abstract, 13 March 1995, see entire reference.	1-3
X	GenBank Database, Accession Number AA343613, ADAMS et al. 'Initial assessment of human gene diversity and expression patterns based on 83 million nulceotides of cDNA sequence', abstract, 21 April 1997, see entire reference.	1, 2

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"B"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 28 OCTOBER 1999	Date of mailing of the international search report 19 NOV 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer CLAIRE M. KAUFMAN Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/18117

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	OHNO et al. *FTZ-F1β*, a novel member of the *Drosophila* nuclear receptor family. Mech. Dev. 1993, Vol. 40, pages 13-24, especially Figure 1.	1-7, 9, 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/18117

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-11,34,35

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/18117

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C07K 14/435, 14/00, 7/06; C12N 5/10, 15/10, 15/11, 15/12, 15/63; C12Q 1/68; A61K 38/16

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

536/23.1, 23.5; 530/300, 350; 435/69.1, 320.1, 325, 252.3, 254.11; 514/2, 12

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-11, 34 and 35, drawn to polynucleotide of clone dj167_2, vector, host cell, method of producing a protein, encoded protein, and method of treating with the protein.
Group II, claim(s) 12-22, 34 and 35, drawn to polynucleotide of clone dj167_19, vector, host cell, method of producing a protein, encoded protein, and method of treating with the protein.
Group III, claim(s) 23-35, drawn to polynucleotide of clone dw665_4, vector, host cell, method of producing a protein, encoded protein, and method of treating with the protein.
The inventions listed as Groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Group I corresponds to the first invention wherein the first product is the polynucleotide and the first method of using is the method of making the protein. The invention also includes the protein made, gene corresponding to the polynucleotide, and method of treating with the protein. Each group does not share the same or corresponding special technical feature because each group is drawn to a different polynucleotide and encoded protein. This Authority therefore considers that the several inventions do not share a special technical feature within the meaning of PCT Rule 13.2 and thus do not relate to a single general inventive concept within the meaning of PCT Rule 13.1